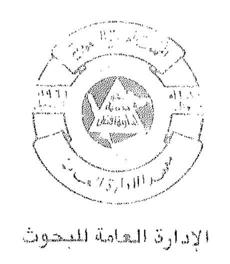


الإدارة العامة للبحوث

الحاسب الآلي والتطبيقات الإحصائية

محمد عثمان البشير   كرم الله علي عبدالرحمن





الإدارة العامة للبحوث

الحاسب الآلي والتطبيقات الإحصائية

بلغة بيسك

محمد عثمان البشير

كرم الله علي عبد الرحمن

معهد الادارة العامة

١٤١١هـ - ١٩٩٠م

المقـــدمـة

العـلاقـة بين الحاسب الآلى والإحصاء ليست فى حاجة لتوضيح ، فهى خاصة ودقيقة. فالإحصاء مجال تطرح فيه مسائل كثيرة ومعقدة، والمبرمجون يلعبون دوراً أساسياً فى حل تلك المسائل، ومن هنا لعب الحاسب دوراً كبيراً فى تطور أساليب التحليل الإحصائى ودقتها. كذلك أصبح للإحصاء فضل كبير فى تطور البرمجة؛ لتميزه بوفرة البيانات التى تحتاج للمعالجة وفق أسس معلومـة، مما جعـل كلاً منهـما مكمـلاً للآخـر خاصة فى مجالات التعليم، والتـدريب، والبحـوث، فأنشئت المؤسسـات المتخصصة فى إعداد البرامج (الحقائب) الإحصائية وتنافست فى تقنية التحليل الإحصائى.

لقـد جاء هذا الكتـاب لتعـزيز تلك المفـاهيم بين البـاحثين والدارسين العرب، ولا يحتاج القارىء لمتابعته إلى غير الإلمام بمبادىء الجبر. وهو يتكون من أحد عشر فصلاً.

موضوع **الفصل الأول** هو «حل مسألة بواسطة الكمبيوتر» بصفة عامة، أما **الفصل الثانى** فقد تخصص فى لغة البيسك المستخدمة فى بقية الفصول. ويبدأ التداخل بين البرمجة والإحصاء فى **الفصل الثالث** الذى تناول مفاهيم تبـويب البيانـات الإحصائية، ووصفها إحصائياً، وبعد حل بعض الأمثلة يدوياً استخدمت لغة البيسك لحل نفس الأمثلة، وهكذا اتبع هذا الأسلوب فى بقية الفصول، إذ نبدأ بتوضيح المفاهيم الإحصائية ومجالات التطبيقات الخاصة بها، وبعد حل بعض التمارين يدوياً يأتى دور البرمجة بلغة بيسك لحل نفس التمارين؛ لكى يكون القارىء قادراً على استخدام الأساليب الإحصائية السليمة وتحليل بياناته بواسطة الحاسب الآلى مستخدماً لغة البيسك.

اختص **الفصل الرابع** بمقاييس النزعة المركزية، وتلاه **الفصل الخامس** لمقاييس التشتت والعزوم. أما **الفصل السادس** فهو الخاص بأهم التوزيعات الإحصائية، ويميل كثيراً لشرح بعض النظريات ليعطى تفسيراً للإحصائيات المستخدمة فى جميع الفصول التالية. هذا ويمكن للباحث عدم التعرض لهذا الفصل إن لم يكن فى حاجة لتلك التفسيرات.

أما **الفصل السابع** فقد عالج موضوع حدود وفترات الثقة، وجاء **الفصل الثامن** مكملاً له بموضوع التطبيقات الخاصة باختبارات الفرضيات المعلمية. هذا ولقد اتضحت لنا بواسطة بعض الزملاء الباحثين فى المجالات الاجتماعية والإدارية أهمية التركيز على الاختبارات اللامعلمية؛ لذلك فقد أفردنا **الفصل التاسع** كاملاً لهذا الموضوع، مع توضيح مجال استخدام كل أسلوب. كذلك اتبعنا نفس الأسلوب فى **الفصل العاشر** عندما تعرضنا لأكثر أنواع الارتباط استخداماً. وأخيراً جاء موضوع الانحدار الخطى فى **الفصل الحادى عشر** لاعتماده على الارتباط.

لقد كان جل همنا هو ربط المواضيع، مع عدم إلزام القارىء بمتابعة ما لا يحتاج إليه سواء فى مجال البرمجة أو الإحصاء، مراعين تنوع حاجات المستفيدين. لذلك فقد قسمنا الجداول الإحصائية إلى ثلاثة أنواع : النوع الأول يخص موضوعاً بعينه، ولذلك جاء مع ذلك الموضوع. أما النوع الثانى فيستخدم فى أكثر من موضوع فى فصل واحد، فجاء فى مؤخرة ذلك الفصل. أما النوع الثالث فيستخدم فى أكثر من فصل، ولذلك جاء فى الملاحق بنهاية الكتاب.

وختاماً يجب ألا تفوتنا هذه الفرصة لتقديم شكرنا للكثير من الزملاء الذين أبدوا ملاحظات قيمة ومفيدة حول بعض مواد الكتاب.

المؤلفان

المحتويات

صفحة

الفصل الأول : خطوات ووسائل حل مسألة بواسطة الكمبيوتر ٩
١ - مقدمة ١١
٢ - لغات البرمجة ١١
٣ - الخوارزميات ١٢
٤ - خريطة سير العمليات ١٢
٥ - أنماط خرائط سير العمليات ١٦
٦ - شبه الجفرة ٢٢
٧ - خوارزميات أساسية ٢٦

الفصل الثانى : مقدمة فى لغة بيسك ٣١
١ - مدخل ٣٣
٢ - المكونات الأساسية لبرنامج بيسك ٣٤
٣ - العمليات الحسابية ٣٨
٤ - أوامر الإدخال ٤٠
٥ - نقل التسلسل والمقارنة ٤٢
٦ - الدوارة ٤٥
٧ - النسق والمصفوفات ٤٨
٨ - الدوال ٥٢
٩ - دوال المبرمج ٥٣
١٠ - عبارات إخراج متقدمة ٥٤
١١ - البرامج الفرعية ٥٨
تمارين ٦١

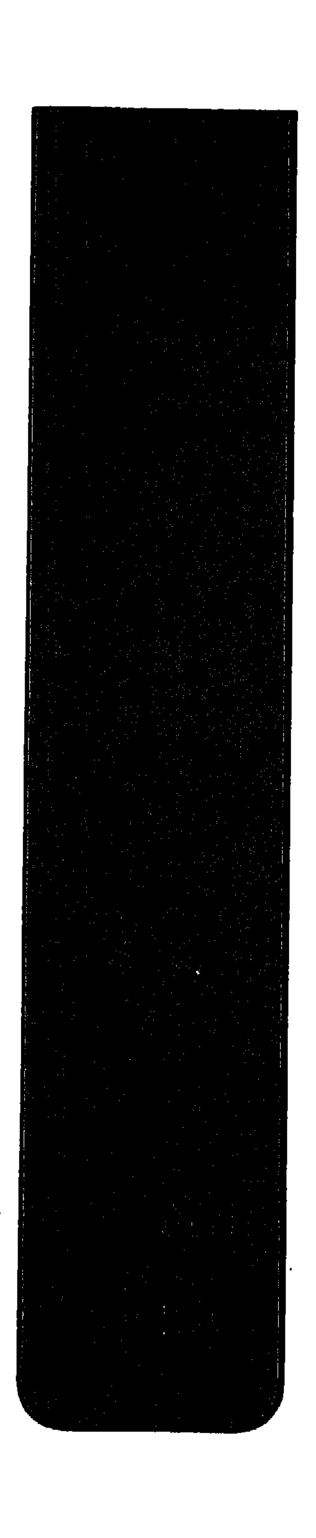
الفصل الثالث : التوزيعات التكرارية لبيانات العينة ٦٩
١ - المقدمة ٧١
٢ - أنواع البيانات ٧٢
٣ - تبويب البيانات الوصفية البسيطة ٧٣
٤ - تبويب البيانات الكمية ٧٦
٥ - التجمع التكرارى ٨٥
٦ - العرض البيانى ٨٩
تمارين ٩٤

	صفحة
الفصل الرابع : مقاييس النزعة المركزية	٩٩
١ - الإحصائية	١٠١
٢ - الوسط الحسابى	١٠٢
٣ - الوسيط	١١٤
٤ - خصائص الوسيط واستخداماته	١١٩
٥ - المنوال	١٢٠
٦ - خصائص المنوال واستخداماته	١٢٣
٧ - العلاقة بين الوسط والوسيط والمنوال	١٢٤
٨ - الوسط الهندسى	١٢٥
٩ - خصائص الوسط الهندسى واستخداماته	١٢٩
١٠ - الوسط التوافقى	١٣٠
١١ - خصائص الوسط التوافقى واستخداماته	١٣٥
١٢ - الربيعات والعشيرات والمئينيات	١٣٦
تمارين	١٤٣
الفصل الخامس : مقاييس التشتت والعزوم	١٤٧
١ - المقدمة	١٤٩
٢ - المدى	١٥٠
٣ - الانحراف الربيعى	١٥٠
٤ - الانحراف المتوسط	١٥٥
٥ - الانحراف المعيارى	١٥٧
٦ - الانحراف المعيارى والمقارنات	١٦٤
٧ - العزوم	١٧١
٨ - الالتواء	١٧٤
٩ - التفرطح	١٨٣
تمارين	١٩١
الفصل السادس : أهم التوزيعات الاحتمالية	١٩٥
١ - المتغير العشوائى	١٩٧
٢ - التوزيع الطبيعى	١٩٨
٣ - التوزيع ذو الحدين	٢٠٠
٤ - توزيع مربع كاى	٢٠٣
٥ - توزيع ف	٢٠٩
٦ - توزيع ت	٢١١

	صفحة
٧ - توزيع الوسط الحسابى للعينة	٢١٣
٨ - توزيع مجموع الوسطين أو الفرق بينهما	٢٢٠
٩ - توزيع نسبة المجتمع	٢٢٣
تمارين	٢٢٧
الفصل السابع : فترات الثقة	٢٢٩
١ - الاستدلال الإحصائى	٢٣١
٢ - فترات الثقة للأوساط	٢٣٢
٣ - فترة الثقة للفرق بين وسطين	٢٣٩
٤ - حدود الثقة للنسب	٢٤٤
٥ - فترات الثقة للتباينات	٢٤٨
تمارين	٢٤٩
الفصل الثامن : تطبيقات اختبارات الفرضيات	٢٥١
١ - تعريف الفرضية والاختبار	٢٥٣
٢ - القرار	٢٥٨
٣ - اختبارات الوسط الحسابى لعينة واحدة	٢٦٠
٤ - اختبارات الفرق بين وسطين من عينتين مستقلتين	٢٦٥
٥ - اختبار الفرق بين وسطين لأزواج متشابهة أو لعينة واحدة	٢٧٢
٦ - اختبار الفرق لأكثر من وسطين	٢٧٥
٧ - اختبارات النسب	٢٧٥
٨ - اختبارات التباين	٢٨٣
تمارين	٢٩٠
الفصل التاسع : تطبيقات الاختبارات غير المعلمية على البيانات الأسمية والتسلسلية	٢٩٥
١ - الفرق بين الاختبارات المعلمية واللامعلمية	٢٩٧
٢ - اختبارات البيانات الاسمية	٢٩٨
٣ - اختبارات البيانات التسلسلية	٣١١
٤ - اختبارات الاستقلال بجداول التوافق	٣٤٤
٥ - تمارين	٣٥٣

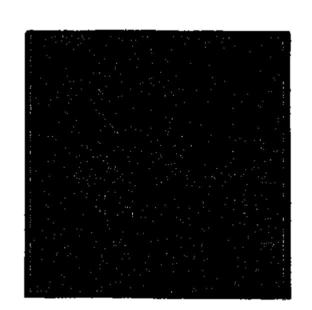
صفحة

الفصل العاشر : الارتباط	٣٥٩
١ - التغاير	٣٦١
٢ - معامل الارتباط الخطى للبيانات النسبية	٣٦٨
٣ - معنوية الارتباط	٣٧٢
٤ - اختبار الفرق بين ارتباطين لعينتين	٣٧٦
٥ - معامل ارتباط الرتب لمتغيرين تسلسليين	٣٧٨
٦ - الارتباط الجزئى	٣٨١
٧ - الارتباط الثنائى التسلسل	٣٨٣
٨ - معامل الارتباط الرباعى للتقسيم الاصطناعى	٣٩٠
٩ - الارتباط بين المتغيرات الاسمية	٣٩١
تمارين	٣٩٥
الفصل الحادى عشر : الانحدار الخطى	٣٩٩
١ - مفهوم الانحدار	٤٠١
٢ - معادلة الانحدار الخطى البسيط	٤٠٢
٣ - خصائص معادلة الانحدار الخطى البسيط	٤٠٨
٤ - انحرافات التقديرات	٤١٣
٥ - الانحدار الثنائى	٤١٤
٦ - الانحدار بالمصفوفات	٤١٧
تمارين	٤٤٩
الملاحق (الجداول الإحصائية)	٤٥٧
المراجع	٤٧٩



خطوات ووسائل حل
مسألة بواسطة الكمبيوتر





خطوات ووسائل حل مسألة بواسطة الكمبيوتر

١ - مقدمة :

لحل أى مسألة بواسطة الكمبيوتر لا بد من استخدام برنامج ما ، سواء أكان هذا البرنامج جاهزاً ومعداً من قبل ، أو كان عليك أن تكتبه بنفسك.

والبرنامج هو سلسلة من الأوامر والتعليمات المرتبطة منطقياً والمكتوبة بإحدى لغات البرمجة ، والتى توجه الكمبيوتر ليؤدى مهام معينة ويصل إلى نتائج محددة.

٢ - لغات البرمجة :

كما أسلفنا الذكر فإن البرنامج يكتب بإحدى لغات البرمجة ، فما هى هذه اللغات؟

قبل الخوض فى ماهية لغات البرمجة دعنا نتعرف على بعض الحقائق الخاصة بالتفاهم بين الإنسان والكمبيوتر. ونبدأ بتقرير حقيقة هامة وهى أن الكمبيوتر لا يفهم إلا لغة الأرقام ، وعليه فكل أجهزة الكمبيوتر الأولى كانت تستخدم ما يعرف بلغات الآلة MACHINE CODE ثم تطورت استخدامات الكمبيوتر وتشعبت ، وأصبحت الحاجة ماسة لكتابة العديد من البرامج لقطاع واسع من التطبيقات ، فكان لا بد من استنباط لغات أكثر سهولة من لغات الآلة ، فكان أن ظهرت لغات المجمع ASSEMBLY والتى لم تحل إلا القليل من مشاكل لغات الآلة . بعد ذلك ظهرت لغات المستوى العالى HIGH LEVEL LANGUAGES والمستخدمة حالياً فى الغالبية العظمى من تطبيقات الكمبيوتر.

«لغات المستوى العالى» أطلق عليها هذا الاسم للتفريق بينها وبين لغات الآلة ولغات المجمع ، والتى تعرف بلغات المستوى البسيط LOW LEVEL LANGUAGES.وقد سميت بهذا الاسم لقربها من مستوى الآلة ، إذ لكل جهاز كمبيوتر لغة الآلة الخاصة به ، وهى مرتبطة بتكوين الدوائر المنطقية الداخلية للجهاز.

لغـات المستوى العالى صممت بحيث تكون سهلة فى التعلم للإسراع فى كتابة البرامج وتعديلها متى ما تطلب الأمر؛ لذلك فهى تكتب بطريقة تشبه إلى حد كبير الكلام الإنجليزى العادى. وهنالك العديد من هذه اللغات إلا أن أكثرها استخداماً هى :

بيسك	BASIC
كوبول	COBOL
فورتران	FORTRAN
ألجول	ALGOL
باسكال	PASCAL
ب ل ١	PL/1

٣ - الخوارزميات[١] ALGORITHMS :

قد تكون عملية كتابة برنامج ما عملية سهلة، ولا تحتاج لكثير من الجهد، إذا كانت المسألة المـراد حلهـا بسيطة وسهلة. أما إذا كانت المسألة معقدة بعض الشيء، فإن هذه العملية تستغـرق الكثـير من الـوقت، وتتـطلب جهـداً إضافياً؛ لذلك كان لا بد من إيجاد وسائل للمسـاعدة فى كتابة البرامج. هذه الوسائل تؤدى إلى تفتيت المسألة إلى عناصر أولية، وإلى خطوات منـطقية تجعـل من السهـولـة كتـابـة الـبرنـامـج. هذه الخـطوات المنـطقية تعرف بالخوارزميات. إذاً فالخوارزمية ما هى إلا خطوات منطقية لحل مسألة ما.

يمكن التعبير عن الخوارزمية بعدة وسائل، وسنتطرق هنا إلى اثنتين من هذه الوسائل وهما :

- خريطة سير العمليات FLOWCHART
- شبه الشفــرة PSEUDOCODE

٤ - خريطة سير العمليات :

خريطة سير العمليات هى رسم بيانى تخطيطى للخطوات التى ينبغى للحاسب أن يتبعها لحل أى مسـألـة. وتستخدم بعض الأشكال لرسم هذه الخرائط، وهنالك العديد من هذه الأشكـال قد تختلف من مؤسسـة لأخرى. وقد جرت عدة محاولات لتوحيد هذه الأشكال

(١) نسبة الى عالم الرياضيات الإسلامى أبى جعفر محمد بن موسى الخوارزمى.

والرموز، من هذه المحاولات ما قام به المعهد الأمريكى الوطنى للمواصفات والمقاييس الذى تبنى الأشكال التالية :

بداية/ نهاية

توصيل من جزء من الخريطة إلى جزء آخر

معالجة/ تشغيل

مدخلات/ مخرجات

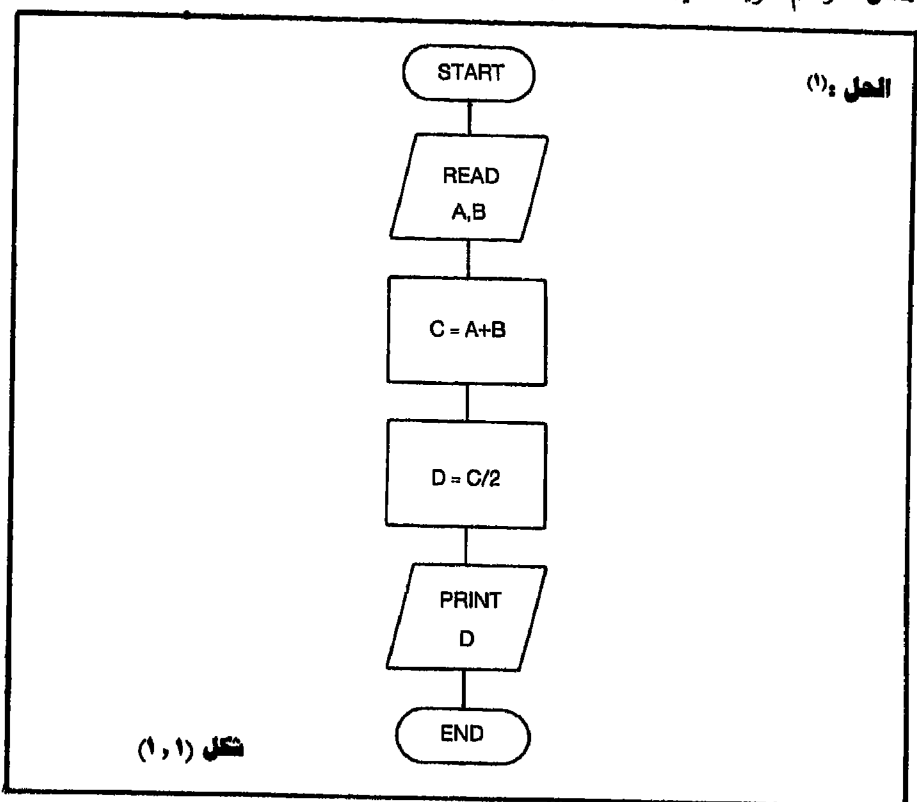
قرار

تجهيز

الأشكال الأساسية المستخدمة فى رسم خرائط سير العمليات

دعنا الآن نأخذ بعض الأمثلة :

مثال : ارسم خريطة سير عمليات لقراءة رقمين وطباعة الوسط الحسابي لهما.



شكل (١ , ١)

الشكل (١ , ١) يمثل أبسط أنواع خرائط سير العمليات، وهو النوع الذى ينساب من أعلى إلى أسفل بدون أى تحويلات فى مساره. وهذا بطبيعة الحال مثال غير عملى إذ قد لا تحتاج المسألة من هذا النوع إلى برنامج كمبيوتر لحلها، إلا أن هنالك ملاحظات ينبغى ذكرها:

١ ـ كل خرائط سير العمليات تبدأ بـ (بداية START) وتنتهى بـ (نهاية END) .

٢ ـ انسياب الخريطة يكون من أعلى إلى أسفل، ما لم تعترضه تحويلات تغير مساره، كما سنرى فيما بعد .

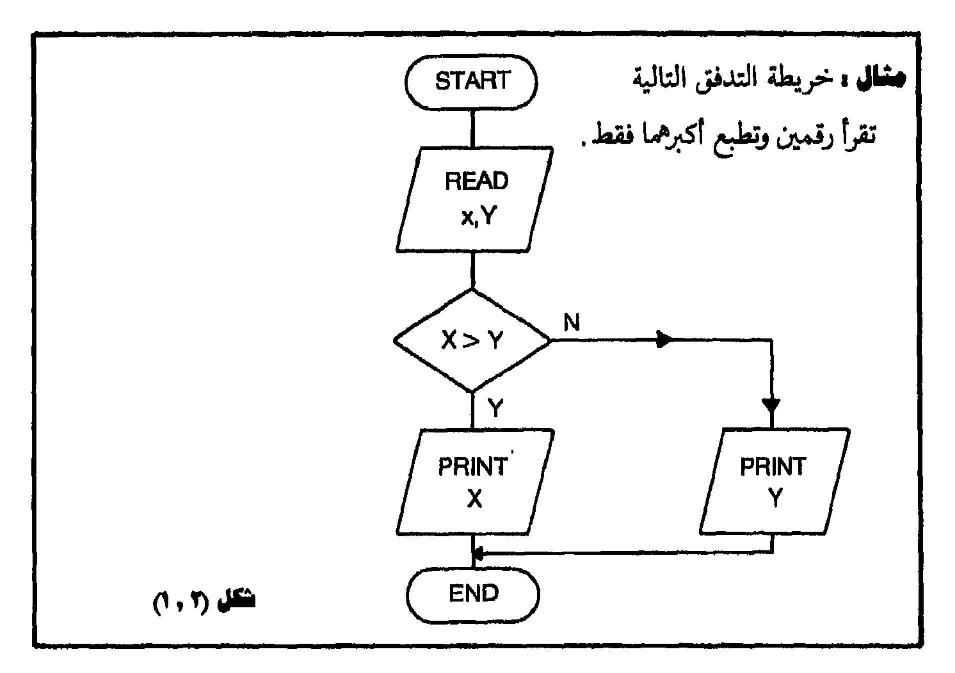
(١) استخدمنا اللغة الإنجليزية فى رسم خرائط العمليات لعدة اعتبارات أهمها تسهيل كتابة البرنامج مباشرة من الخريطة إلى لغة بيسك.

٣ ـ كل خطوة (شكل) من خطوات الخريطة ينبغى أن تكون متصلة من جانبين، لتوضيح الخطوة السابقة عليها والخطوة التى تليها. (ما عدا بالطبع البداية والنهاية).

تحويل المسار BRANCHING :

هذا النمط من خرائط سير العمليات يختلف عن الأول فى أن انسيابه يتحول فى مرحلة من المراحل إلى أحد المسارات أو الآخر، اعتماداً على نتيجة قرار معين، لذلك فهو يستخدم شكل القرار DECISION ، والذى يكون نتيجته نعم أو لا . وهنا تستخدم الرموز والإشارات التى تختبر العلاقة بين قيمتين RELATIONAL والتى نوردها فيما يلى :

الاشارة	**مدلولها**
=	يساوى
>	أكبر من
<	أصغر من
>=	أكبر من أو يساوى
<=	أصغر من أو يساوى
<>	لا يساوى



شكل (٢ ، ١)

الدوارة LOOP :

المثـال (١ , ١) كان يمثـل خريطة سير عمليات لقراءة رقمين فقط، واستخراج وطباعة الوسط الحسابى لهما. وكما ذكرنا فإن ذلك المثال ليس بعملى، فأنت دائماً تستخدم الكمبيوتر لحل المسائل المعقدة الكبيرة، والتى تحتاج لعمليات كثيرة ومتكررة. فمثلاً إذا أردنا أن نوجد المتوسط الحسابى لدرجات ٣٠ طالباً فى امتحان معين، فسيكون من العسير إعطاء كل درجة من هذه الدرجات رمزاً مثل C, B, A... وستكون المسألة أكثر عسراً إذا زاد عدد الدرجات أكثر.

هنا نستغل خاصية مفيدة جداً من خصائص الكمبيوتر، وهى قدرته على تكرار عملية معينة أو مجموعة عمليات، أى عدد من المرات، ويسمى هذا التكرار بالدوارة.

مثال : المثال التالى ـ الشكل ٣ , ١ ـ يقوم بقراءة درجات ٣٠ طالباً فى امتحان معين واحتساب متوسط الدرجات وطباعته.

لاحظ الآتى فى هذا المثال :

- استخدمنا رمز التهيئة لتنظيف حقلى المجموع T والعداد C وذلك بوضع القيمة صفر فيهما كقيمة ابتدائية، إذ أننا استخدمنا الأول لعملية الجمع التراكمى للدرجات، والثانى لعد الدرجات نفسها حتى نحدد نهاية الدوارة.
- استخدمنا متغيراً واحداً هو X قرأنا فيه كل القيم، وهو بهذا يأخذ قيمة متغيرة فى كل دورة. وفى كل دورة فإننا نضيف قيمة X إلى المجموع السابق T و نضيف 1 إلى العداد C .
- لتحديد نهاية الدوارة فقد استخدمنا رمز القرار لمعرفة إذا كانت الأرقام كلها قد قرئت، إذ أننا نسأل إذا كان العداد C أقل من ٣٠ ـ وهى قيمته النهائية ـ فإن كانت الاجابة بلا فإننا نرجع لبداية الدوارة، وإلا فإننا نحسب الوسط الحسابى ونطبعه وننهى الخريطة.

٥ ـ أنماط خرائط سير العمليات :

كما عرضنا آنفا فهنالك عدة طرق لتمثيل الخوارزميات عن طريق خرائط سير العمليات، إلا أن المتحمسين للطرق الحديثة فى البرمجة يقررون أن أى خوارزميـة يمكن تمثيلها بإحدى الطرق الثلاث التالية :

١ ـ منطق تسلسلى SEQUENCE

٢ ـ منطق اختيار SELECTION

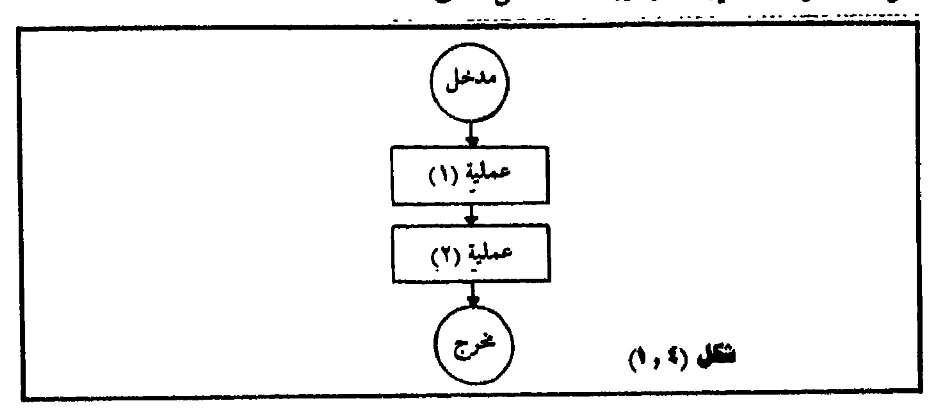
٣ ـ منطق تكرار ITERATION

START
T = ϕ
C = ϕ
READ X
T = T + X
C = C + 1
C < 3ϕ ?
Y
N
V = T/3ϕ
PRINT V
END

شكل (١،٣)

١ - المنطق التسلسلي :

وهو الذى يتم فيه تنفيذ العمليات حسب ترتيبها من أعلى إلى أسفل، دون أن تعترضه شروط أو قرارات تغير مساره ويتخذ الشكل التالى :



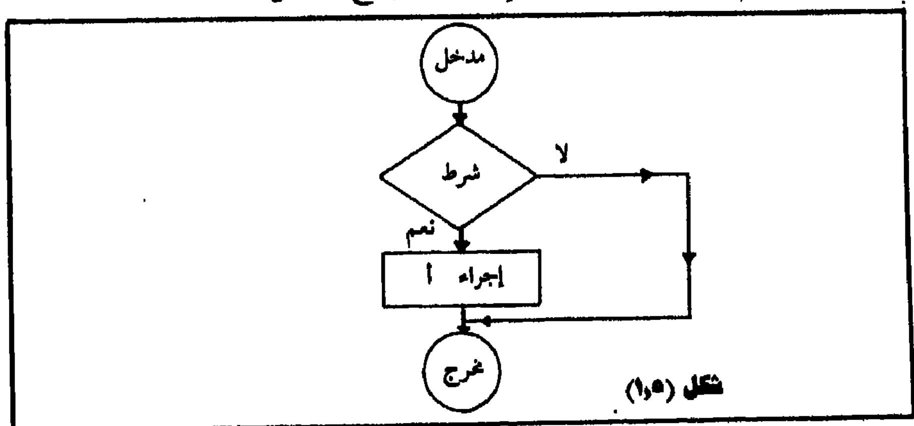
شكل (٤ , ١)

٢ - منطق الاختيار :

وهذا يختلف عن سابقه فى أن هنالك شرطاً معيناً وكنتيجة لهذا الشرط يتغير المسار وهو يستخدم (إذا) الشرطية أو IF وله نوعان أساسيان :

١- مفرد البديل[1] SINGLE ALTERNATIVE :

حيث يكون هنالك شرط، فإن تحقق نفذ إجراء معين (والذى قد يحتوى على عدة خطوات)، وإن لم يتحقق لا ينفذ ذلك الإجراء. ويوضح الشكل ٥ , ١ ذلك.

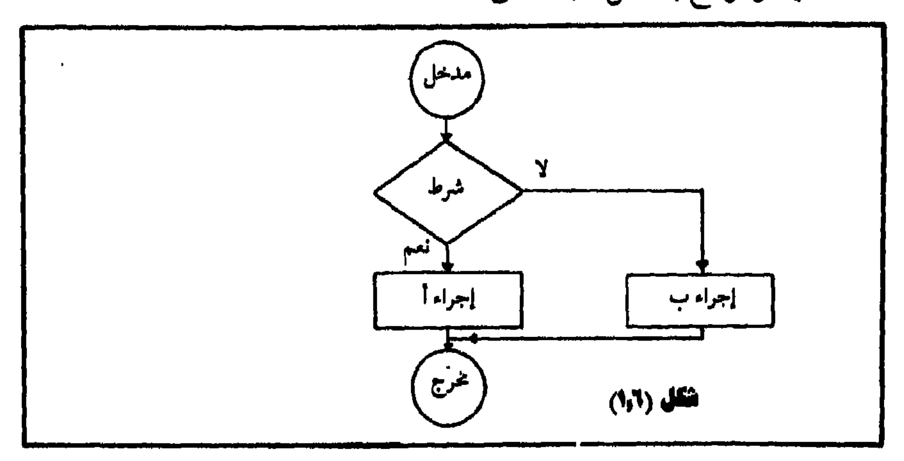


شكل (٥,١)

(1) Lipschutz, Seymour, Essential Computer Mathematics, Mcgraw- Hill Book Company, 1982, P 109.

٢ - مزدوج البديل Double Alternative :

حيث يكـون هنالك شرط، فإن تحقق نفذ إجراء معين، وإن لم يتحقق نفذ إجراء آخر مختلف كما هو موضح بالشكل ١,٦ التالى :



شكل (١,٦)

٣ - منطق التكرار :

وهذا يختص بالدوارة أو تنفيذ عملية معينة - أو مجموعة عمليات - عدداً من المرات وهو يتخذ ثلاثة أشكال :

الشكل الأول :

وهو الذى يستخدم مؤشراً معيناً لتحديد عدد المرات التى ينفذ فيها الإجراء أو مجموعة الإجراءات.

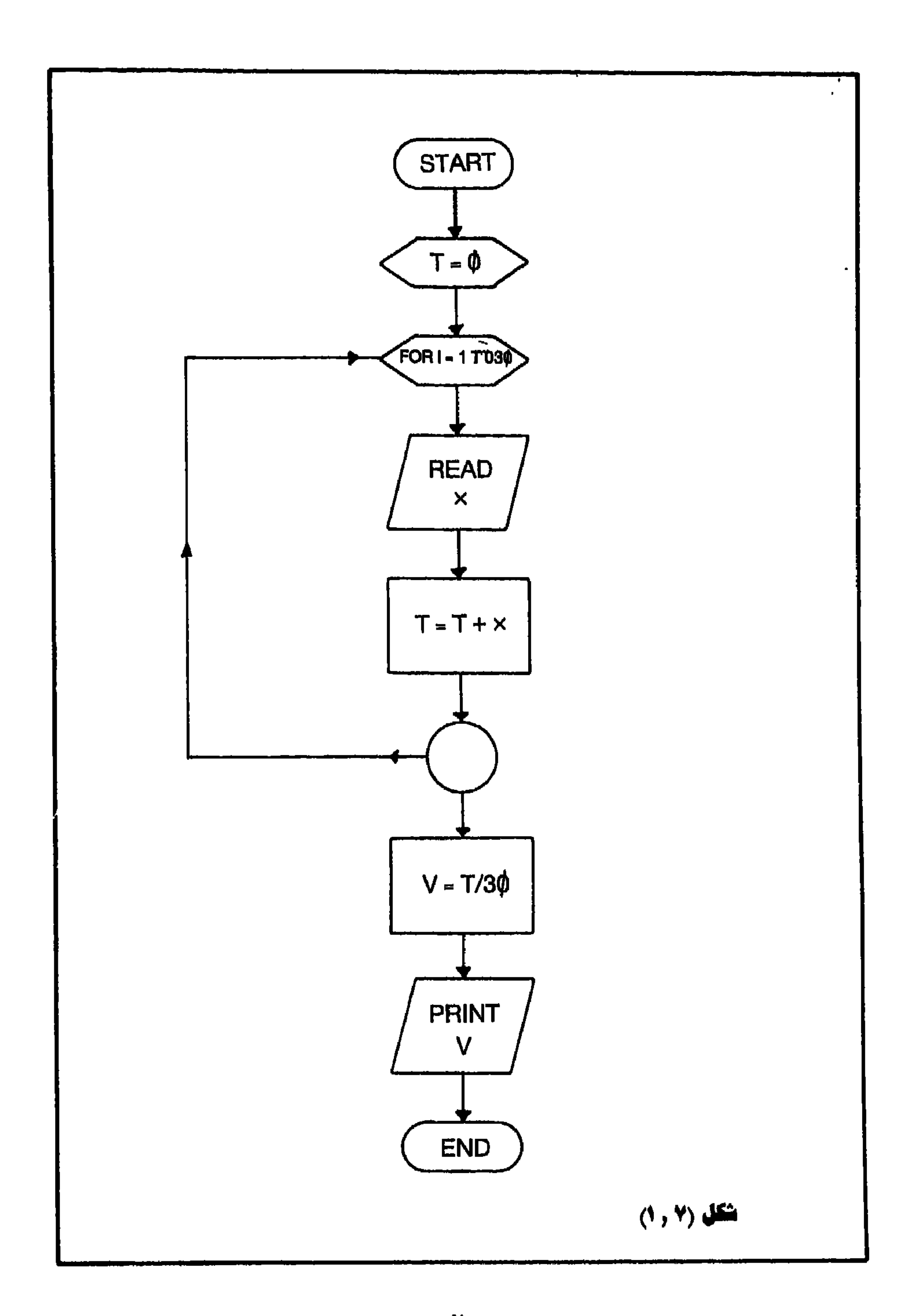
وعدد المرات يتحدد بأن المؤشر يأخذ قيمة ابتدائية وقيمة نهائية وإضافة. وله عدة تسميات حسب لغة البرمجة المستخدمة فهو يستخدم عبارة DO مثل :

```
DO R = 1 TO N BY 1
```

أو فى لغة بيسك يأخذ عبارة FOR ...NEXT مثل :

```
FOR J = 1 TO N STEP 1
```

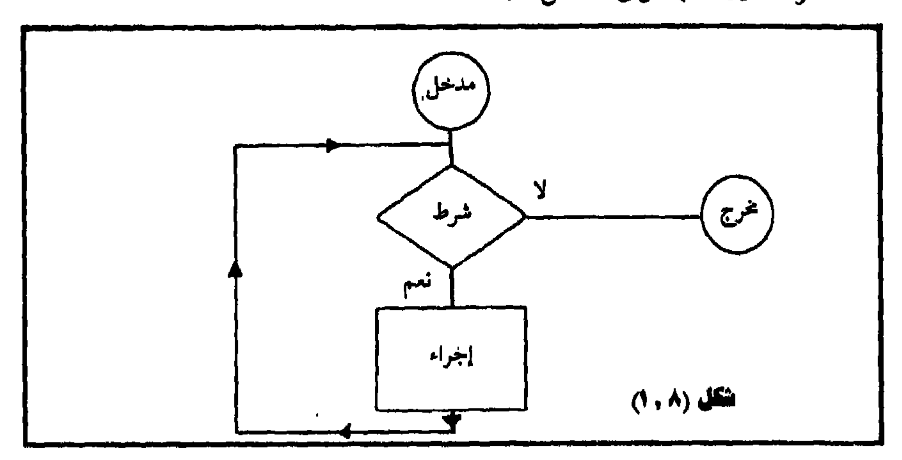
والشكل ١,٧ يوضح هذا النوع من التكرار حيث يبين حل مسألة لقراءة ٣٠ رقماً وطباعة متوسطها.



شكل (٧ , ١)

الشكل الثاني :

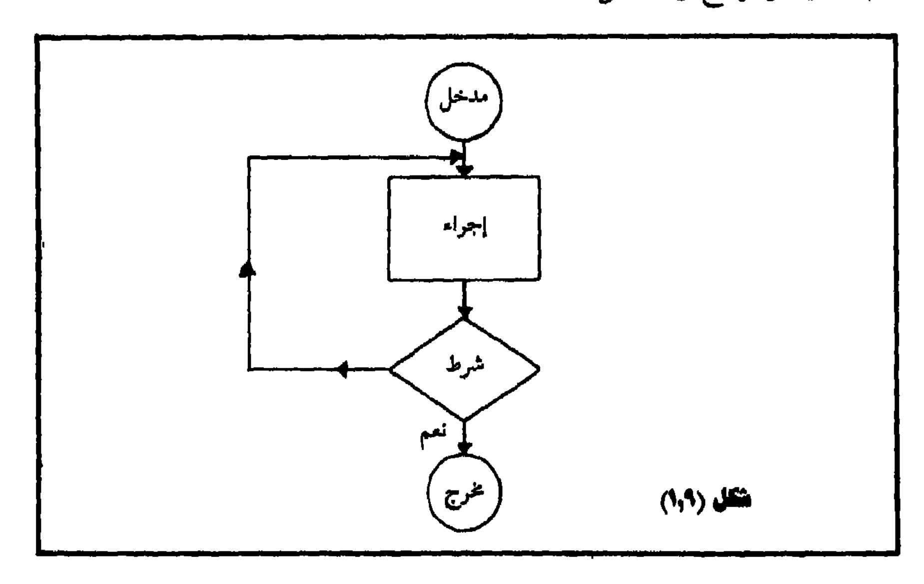
وهـذا يعرف بتركيبة DO WHILE وهو أن ينفذ إجراء معين عدداً من المرات اعتماداً على صحة شرط معين، كما هو في الشكل ١,٨ .



شكل (١,٨)

الشكل الثالث :

وهذا يستخدم تركيبة DO UNTIL وهو أن تستمر في تنفيذ إجراء معين إلى أن يتحقق شرط معين، كما هو موضح في الشكل ١,٩ .



شكل (١,٩)

٦ ـ شبه الجفرة :

فى العديد من الحالات قد لا تفى خرائط سير العمليات بالغرض المطلوب، وهو التعبير عن الخوارزمية، وجعل كتابة البرنامج أمراً سهلاً. فخريطة سير العمليات قد لا تكون كافية، وربما تكون صعبة الفهم لبعض المسائل المعقدة. وعليه فهنالك طريقة قد تكون مكملة لخرائط سير العمليات، أو قد تغنى عنها فى كثير من الحالات وهى شبه الجفرة.

ميزة شبه الجفرة أنها تكتب بطريقة تشبه الكلام العادى، لذلك فهى سهلة فى الفهم والكتابة.

دعنا نرى الآن كيف يمكننا تمثيل الأشكال الثلاثة التى تطرقنا لها فى خرائط سير العمليات بواسطة شبه الجفرة :

المنطق التسلسلى :

وهو الذى يتخذ الشكل التالى :

.
.
.
عملية أ
عملية ب
.
.
.

مثال :

```
READ SPEED, TIME
DISTANCE = SPEED X TIME
PRINT SPEED, TIME, DISTANCE
END
```

منطق الاختيار :

وهو الذى تستخدم فيه أداة المقارنة IF ويبدأ بـ IF وينتهى بـ END IF وله شكلان كما رأينا من قبل.

مفرد البديل :

ويتخذ الشكل التالى :

```
IF Condition THEN
   Procedure A
END IF
```

مثال :

```
READ DISTANCE, TIME
SPEED = DISTANCE / TIME
IF SPEED >1ØØ THEN
   PRINT "HIGH SPEED"
END IF
PRINT SPEED, DISTANCE, TIME
END
```

لاحظ كيفية كتابة العمليات بين IF و END IF حيث تكون إلى الداخل قليلاً، وهو إجراء متعارف عليه لتسهيل قراءة وفهم شبه الجفرة .

مزدوج البديل :

وهو الذى يستخدم فيه تعبير IF... THEN... ELSE ويتخذ الشكل التالى :

```
IF Condition THEN
   Procedure A
ELSE
   PRocedure B
END IF
```

مثال :

```
READ DISTANCE, TIME
SPEED = DISTANCE / TIME
IF SPEED>1ØØ THEN
     PRINT "HIGH SPEED"
ELSE
     PRINT "LOW SPEED"
END IF
PRINT DISTANCE, TIME, SPEED
END
```

منطق التكرار :

وهو كما رأينا يتخذ ثلاثة أشكال سنستعرضها فيما يلى :

- الشكل الأول :

وهو كما أسلفنا يستخدم تركيبة :

DO I = 1 TO N BY 1

حيث I هو المتغير الذى يتحكم فى الدوارة

N هو المتغير الذى يمثل القيمة النهائية للدوارة . وعليه فكل التركيبة تعنى أن ننفذ الإجراء أو الإجراءات التالية بقيمة I تساوى 1 إلى أن تصبح قيمة I تساوى N مع زيادة 1 إلى I فى كل دورة .

مثال :

سنعيد كتابة نفس المثال الذى مثلناه فى حديثنا عن خرائط سير العمليات ـ المثال التالى ـ باستخدام شبه الجفرة .

```
T = ∅
READ N
┌ DO I = 1 TO N BY 1
│     READ X
│     T = T + X
└ END DO
V= T/N
PRINT V
END
```

- الشكل الثانى :

وهذا يكتب باستخدام عبارة DO WHILE ويتخذ الشكل العام التالى :

```
DO WHILE Condition
     Procedure
END DO
```

مثال :

المثال التالى يمثل شبه جفرة لقراءة مجموعة من الأرقام يرمز لعددها بالمتغير N وهو مجهول وطباعة الأرقام واحتساب وطباعة مجموعها والوسط الحسابى لها، علماً بأن الخروج من الدوارة يتم حين يكون أحد الأرقام صفراً.

```
      N = ∅
      S = ∅
┌──── DOWHILE  X ≠ ∅
│         READ  X
│         PRINT  X
│         N = N+1
│         S = S + X
└──── END DO
      V = S/N
      PRINT S,V
      END
```

ـ الشكل الثالث :

وهذا الشكل يستخدم عبارة DO UNTIL ويتخذ الصورة التالية :

```
DOUNTIL Condition
     Procedure
END DO
```

مثال :

سنعيد كتابة نفس المثال السابق باستخدام تركيبة DO UNTIL

```
N = ϕ
S = ϕ
DOUNTIL  X = ϕ
    READ  X
    PRINT  X
    N = N + 1
    S = S + X
END DO
V = S/N
PRINT  S,V
END
```

٧ ـ خوارزميات أساسية :

سنحاول فيما يلى أن نتعرض لأهم الخوارزميات ، التى تعتبر أساسية لدراستنا هذه بشكل عام ، ولهذا الفصل بشكل خاص . والعمليات التى سنتعرض لها الآن تعتبر من أكثر العمليات استخداماً ، وستصادفنا كثيراً طوال دراستنا .

١ ـ تبديل قيمتى متغيرين :

كثيراً ما تصادفنا حالات نحتاج فيها إلى تبديل قيمتى متغيرين أثناء معالجتنا لبعض البيانات . أكثر التطبيقات التى نحتاج فيها لهذه العملية هى تطبيقات الفرز.

مثال :

إذا كان لدينا المتغير A وقيمته 27 والمتغير B وقيمته 65 والمطلوب تبديل القيمتين ، بحيث تصبح قيمة A هى 65 وقيمة B هى 27 .

الحل :

الوضع الحالى للمتغيرين

A	B
27	65

الوضع المطلوب

A B

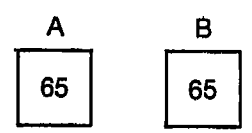
65 27

قد يتبادر الى الذهن مباشرة أن الحل هو كالآتى :

A = B

B = A

وهو بطبيعة الحال حل خاطىء إذ سيؤدى إلى النتيجة التالية :



حيث سيأخـذ المتغيران نفس القيمة 65 وهى قيمة B الأصلية. وذلك لأن عبارة A = B ستأخذ قيمة B وهى 65 وتضعها فى A وبالتالى فقيمة A السابقة ستضيع ولا مجال لاستردادها. أما العبارة الثانية B = A فلن تكون ذات قيمة بعد ذلك.

الحل الصحيح هو أن نحفظ قيمة A الأصلية فى مكان ما مؤقتاً حتى لا تضيع، ومن ثم وضعها فى B وعليه يكون توصيف الخوارزمية كالآتى :

١ ـ احفظ فى C قيمة A الأصلية.
٢ ـ ضع فى A قيمة B الأصلية.
٣ ـ ضع فى B قيمة A الأصلية والمحفوظة فى C .

ويمكن تمثيل هذه الخوارزمية بشبه الجفرة التالية :

C = A

A = B

B = C

٢ ـ العد COUNTING :

من العمليات التى نحتاج إليها كثيراً : عملية العد، فقد نحتاج أن نعد حالات معينة من مجموعة من الحالات. كأمثلة على ذلك : عدد الأشخاص الذين تقل أعمارهم عن ٢٠ من مجموعة من سكان مدينة ما، الأشخاص الذين تزيد رواتبهم على ١٠,٠٠٠ دولار من موظفى مؤسسة ما، الطلاب الذين تقل درجاتهم عن ٦٠ فى اختبار ما، وهكذا.

مثال :

لدينا درجات N من الطلبة فى اختبار ما، والمطلوب عدد الطلاب الذين تقل درجاتهم عن ٦٠.

الحل :

لابد أن يكون لدينا متغير معين نستخدمه كعداد وكلما وجدنا درجة أقل من ٦٠ أضفنا واحداً إلى هذا العداد. لكن السؤال هو: كيف تتم عملية الإضافة هذه؟

تتم هذه العملية بالصورة التالية والتى رأيناها كثيراً عند حديثنا عن شبه الجفرة :

العداد الجديد = العداد السابق + 1

حيث إن العداد الجديد والعداد السابق يتم تمثيلهما بمتغير واحد مثل C ، وهذا معناه أن القيمة الجديدة للعداد تساوى قيمته السابقة مضافاً إليها واحد، كالآتى :

$$C = C + 1$$

وهذه الصورة ستصادفنا كثيراً فى مجال معالجة البيانات، حيث تدخل قيمة المتغير السابقة فى تحديد قيمته الجديدة. فقط هنالك قاعدة ينبغى مراعاتها وهى أنه حين يوجد متغير ما على يمين ويسار علامة = فى نفس العبارة، فلا بد من إعطاء هذا المتغير قيمة ابتدائية. وفى هذه الحالة ـ العد ـ ومعظم الحالات فإن القيمة الابتدائية لهذا المتغير تكون صفراً.

توصيف الخوارزمية :

١ ـ اجعل القيمة الابتدائية للعداد صفراً.

٢ ـ اقرأ عدد القيم N.

٣ - اقرأ إحدى الدرجات.
٤ - إذا كانت الدرجة أقل من ٦٠ فأضف 1 إلى العداد.
٥ - كرر الخطوات ٣ - ٤ إلى أن تقرأ كل القيم N
٦ - اطبع عدد الدرجات أقل من ٦٠.

وهذه الخوارزمية يمكن تمثيلها بشبه الجفرة التالية :

```
C = Ø
READ N
DO I = 1   TO N   BY 1
    READ X
    IF X< 60 THE N
      C = C + 1
    END IF
END DO
PRINT C
```

٣ - جمع عدد من الأرقام SUMMATION :

من العمليات الأساسية والمستخدمة بكثرة هو استخدام الكمبيوتر لجمع مجموعة من الأرقام، وفى مجال دراستنا هذه على وجه الخصوص نجد أن هذه العملية مستخدمة فى الغالبية العظمى من المقاييس الإحصائية.

مثال :

من أهم وأبسط المقاييس الإحصائية هو الوسط الحسابى، وقد رأينا فى مجال حديثنا عن شبه الجفرة كيفية احتساب الوسط الحسابى لمجموعة من الأرقام، وقد رأينا كذلك أنه لكى تحسب الوسط الحسابى فلا بد من جمع كل القيم، ومن ثم قسمتها على عددها.

توصيف الخوارزمية :

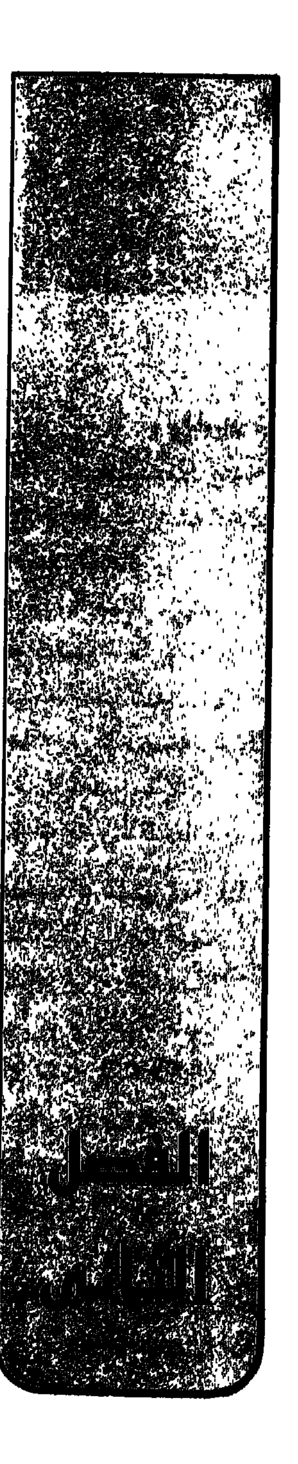
١ - اجعل القيمة الابتدائية للمجموع صفراً.
٢ - اقرأ عدد القيم N .
٣ - اقرأ إحدى القيم.

٤ ـ أضف القيمة إلى المجموع .
٥ ـ كرر الخطوتين ٣ ـ ٤ إلى أن تقرأ كل القيم N .
٦ ـ اطبع مجموع القيم .

وقد رأينا فى حديثنا عن شبه الجفرة كيفية تمثيلها بعدة طرق، ونعيد كتابتها هنا بإحدى هذه الطرق :

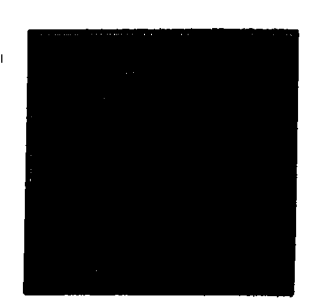
```
T = ∅
READ   N
DO I = 1  TO N  BY 1
   READ   X
   T   = T + X
END DO
PRINT   T
```

لاحظ أننا استخدمنا نفس التعبير T = T + X والذى تحدثنا عنه فى معرض حديثنا عن العد حيث يعنى أن قيمة المجموع الجديد T تساوى قيمته السابقة مضافاً إليها القيمة التى قرئت مؤخراً X .



مقدمة فى
لغة بيسك





مقدمة فى لغة بيسك

١ ـ مدخل :

تعتبر لغة بيسك من أكثر اللغات ذات المستوى العالى شهرة واستخداماً، واسم اللغة BASIC مشتق من العبارة الإنجليزية :

Beginners' ALL-purpose Symbolic Instruction Code

والتى تعنى : اللغة الرمزية المتعددة الأغراض للمبتدئين.

تم تطوير لغة بيسك حوالى عام ١٩٦٣ فى كلية دارتموث لتسهيل تعليم منطق البرمجة للطلاب، وقد استخدمت من قبل الطلاب فى التخصصات المختلفة، ومنذ ذلك الوقت واللغة تكتسب كل يوم شعبية جديدة. وقد أعطى انتشار أجهزة الكمبيوتر الشخصى فى السنوات الأخيرة دفعة كبيرة للغة، إذ أنها توجد ـ تقريباً ـ فى كل هذه الأجهزة، حتى الآلات الحاسبة القابلة للبرمجة بها إمكانات برمجة بلغة بيسك.

تمتاز لغة بيسك بالعديد من المزايا التى أكسبتها كل هذه الشعبية، نذكر منها :

١ ـ سهلة التعلم، إذ تشبه عباراتها إلى حد كبير المعادلات الجبرية العادية، كما أنها تستخدم كلمات إنجليزية مألوفة، مثل READ و PRINT و STOP .

٢ ـ سهلة التنفيذ، إذ أنها مثالية فى الأجهزة التى تستخدم نظام مشاركة الوقت TIME SHARING حيث يجلس المستخدم أمام نهائية مرتبطة بجهاز حاسب كبير، ويكتب البرنامج ويترجمه الحاسب وينفذه فى نفس الوقت.

٣ ـ تتمتع لغة بيسك بمقدرات رياضية عالية إذ بها دوال FUNCTIONS جاهزة مثل الجذر التربيعى واللوغاريتم، ودوال حساب المثلثات وغيرها.

هنالك العديد من نسخ VERSIONS بيسك المطورة، إلا أن المعهد الأمريكى للمواصفات قد أقر بيسك المعيارية ANS BASIC والتى هى بيسك الأساسية دون إضافات، والتى سوف نتبعها فى هذا الفصل.

٢ ـ المكونات الأساسية لبرنامج بيسك :

نكتب البرامج بلغة بيسك سطراً فسطراً، وكل سطر يعبر عن أمر معين أو تعليمة لأداء فعل معين. كل سطر من البرنامج يبدأ برقم، وكل رقم سطر لا بد أن يكون أكبر من رقم السطر الذى قبله، وليست هنالك قاعدة أخرى لترقيم الأسطر، فكل مبرمج حر فى أن يختار التسلسل الذى يناسبه على أن تراعى القاعدة، وهى أن يكون كل سطر أعلى من السطر الذى قبله، على أنه يفضل دائماً أن يكون هنالك فراغات بين أرقام الأسطر تحسباً لإدخال سطر جديد ـ أو عدة أسطر ـ بين سطرين من الأسطر الأصلية.

وتختلف البرامج فى حجمها الذى قد يكون سطراً واحداً فقط هو عبارة END والتى تدل على نهاية البرنامج، ولا بد من وجودها فى كل برنامج. البرنامج التالى برنامج صحيح من الناحية القاعدية للغة :

```
10   END
```

لكنه من الناحية العملية لا يقدم شيئاً للمستفيد الذى يود أن يستخدم الحاسب ليحل مشكلة ما.

دعنا الآن نكتب أول برنامج بلغة بيسك.

مثال (١ , ٢)

تأمل هذا البرنامج :

```
10 PRINT 'THIS IS THE FIRST PROGRAM'
20 END
```

يتكون هذا من تعليمتين هما رقم 10 ورقم 20 وبما أننا ذكرنا أن تعليمة END لا بد من وجودها فى كل برنامج، إذن فالبرنامج يتكون من تعليمة واحدة فقط هى رقم 10 .

عند تنفيذ هذا البرنامج سنرى الآتى :

```
THIS IS THE FIRST PROGRAM
```

الذى فعله الكمبيوتر عند تنفيذ البرنامج هو أنه نفذ التعليمة فى السطر 10 وهى تعليمة PRINT وقد أخرج على الشاشة العبارة التى بعد كلمة PRINT .

تعليمة الاخراج LOOP :

تعليمة PRINT من أهم تعليمات لغة بيسك، فهى الصلة بين المستخدم والجهاز، فكل الرسائل والنتائج التى يخرجها الجهاز يتم توصيلها للمستخدم عن طريق تعليمة PRINT .

تستخدم تعليمة PRINT لإخراج العبارات الثابتة كما رأينا فى المثال السابق، وفى هذه الحالة لا بد من وضع العبارة التى نريد إظهارها بين علامتى ' ' (فى بعض الأجهزة تستخدم علامة " ").

أمثلة :

```
10  PRINT 'MY NAME IS AHMED'
20  PRINT 'ENTER YOUR AGE'
30  PRINT 'PLEASE ENTER TWO NUMBERS'
70  PRINT 'THE ANSWER = '
```

عند تنفيذ الحاسب لأى من التعليمات السابقة نجد أن الجهاز يظهر ما هو مكتوب بين علامتى ' ' كاملاً ودون تصرف.

كذلك يمكن لتعليمة PRINT أن تخرج محتويات حقول متغيرة، وسنتطرق لهذا لدى حديثنا عن تعليمة الإسناد LET . أما الآن فدعنا نتوقف قليلاً لإلقاء نظرة على أنواع البيانات.

الثوابت والمتغيرات :

الثابت هو قيمة ثابتة لا يطرأ عليها أى تغيير، وهى نوعان : ثوابت رقمية ، وهى الأرقام عموماً موجبة أو سالبة، صحيحة أو عشرية، وثوابت حرفية، وهى أى مجموعة من الحروف أو/و الأرقام وأى علامات خاصة تكون محصورة بين علامتى' ' كما رأينا فى المثال السابق.

أما المتغيرات فهى الحقول التى تحمل قيماً متغيرة تتغير قيمها بتغير المحتوى الذى تحمله. وهى كذلك نوعان : متغيرات رقمية، ومتغيرات حرفية.

المتغيرات الرقمية هى التى تحمل بيانات رقمية، مثل : درجات الحرارة، الارتفاعات، السرعات، الأطوال، عدد أفراد الأسرة، الأعمار، الرواتب وغيرها.

أما المتغيرات الحرفية فهى الحقول التى تحمل قيماً حرفية، مثل : الأسماء عموماً، العناوين، رموز قطع الغيار، وصف الأشياء، وهكذا.

للتعبير عن المتغير الرقمى فى برنامج بيسك فإننا ربما نستخدم حرفاً واحداً أو حرفاً واحداً وبعده رقم واحد [1]، فمثلاً هذه أمثلة لمتغيرات رقمية صحيحة :

X

J5

R2

A

وللتعبير عن متغير حرفى فإننا ربما نستخدم أيضاً حرفاً واحداً فقط بعده علامة الدولار $ للتفريق بين المتغير الرقمى والمتغير الحرفى . هذه أمثلة لمتغيرات حرفية :

M $

A $

Y $

G $

عبارة الملاحظات REM :

كما سبق وذكرنا فلغة بيسك لغة رمزية، لذلك فقد يحتاج البرنامج لبعض الشرح لتتبعه وفهمه ولهذا الغرض تستخدم تعليمة REM .

عبارة REM تستخدم لكتابة الملاحظات فى البرنامج . هذه الملاحظات تؤدى إلى توثيق البرنامج، مما يساعد على فهمه من قبل أى شخص يقرؤه ، وبالتالى تساعد على تتبع البرنامج وتعديله إذا لزم الأمر.

تكتب عبارة REM فى أى مكان بالبرنامج، وعندما يجد مترجم اللغة هذه العبارة فى أى مكان، فإنه لا ينظر إلى ما بعدها ويعتبره خاصاً بتوثيق البرنامج .

أمثلة

```
10 REM PROGRAM NO .1
30 REM THIS PROGRAM CALCULATES THE AVERAGE
80 INPUT N REM NO. OF OBSERVATIONS
```

(١) بعض الأنظمة تستخدم أكثر من حرف.

تعليمة الاسناد LET :

البيانات التى تطرقنا لها آنفاً لا بد من وسيلة لتوصيلها للبرنامج . يتم هذا عن طريق تعليمة LET وهى تعليمة أساسية فى لغة بيسك .

تأمل هذه العبارة :

```
20 LET R = 64
```

هذه العبارة معناها أن المتغير الرقمى R ـ بعد تنفيذ هذه التعليمة ـ سيأخذ القيمة على يمين علامة الإسناد = وهى 64 . وينبغى هنا التفريق بين علامة الإسناد = وعلامة يساوى الحسابية . علامة = هنا لا تعنى يساوى الحسابية ، بل هى دليل على إسناد القيمة التى على يمين العلامة إلى المتغير الذى على يسارها . ولا بد من وجود متغير واحد فقط على يسار علامة الإسناد = ، أما على يمينها فيمكن أن نجد واحداً من ثلاثة أشكال :

(١) ثابت مثل	10 LET = 916
أو (٢) متغير آخر مثل	30 LET N = M
أو (٣) تعبير حسابى مثل	70 LET T = F − X + 7

ففى (١) نجد أن المتغير F قد أخذ القيمة 916 ، وفى (٢) نجد أن المتغير N قد أخذ نفس قيمة المتغير M . أما فى (٣) فإن T قد أخذ قيمة التعبير F − x + 7 .

هذا بالنسبة للمتغيرات الرقمية ، أما الإسناد للمتغيرات الحرفية فيتم كذلك بتعليمة LET لكن فى حالة المتغيرات الحرفية لا بد من وجود القيمة المراد إسنادها بين علامتى ' ' كما أوضحنا من قبل .

أمثلة :

```
30 LET A$ = 'PROGRAM 5'
40 LET X$ = 'THE AVERAGE IS'
60 LET N$ = 'MOHAMED ALI'
```

كذلك يمكن للمتغير الحرفى أن يأخذ قيمة متغير حرفى آخر.

أمثلة :

```
10 LET B$ = Y$
50 LET C$ = D$
```

من الجـدير بالذكر أن كلمة LET نفسها يمكن عدم كتابتها مع احتفاظ التعليمة بنفس المعنى، فمثلاً العبارتان :

```
10 LET E = T
10 E = T
```

تؤديان نفس الغرض ، وكذلك العبارتان :

```
60 LET X$ = 'ALI'
60 X$ = 'ALI'
```

تؤديان الغرض نفسه.

٣ ـ العمليات الحسابية :

يتم التعبير عن أى عملية حسابية بطريقة قريبة من الشكل الحسابى العادى، وذلك باستخدام العلامات ـ الإشارات ـ التالية :

العملية	الإشارة	مثال
الجمع	+	X+Y
الطرح	-	X-Y
الضرب	*	X*Y
القسمة	/	X/Y
الرفع للقوة	**	X**Y

وتستخدم تعليمة LET لإجراء هذه العمليات ، ويمكن جمع أكثر من عملية فى تعليمة LET واحدة :

أمثلة :

```
20 LET X = 9 + 5
30 D = A + F
40 P = M - 2 + H
50 LET A = N*R - 1
60 L= E/6*Y** 2
```

لمعرفة ناتج أى تعبير حسابى يحتوى على أكثر من عملية ، فإنك تبدأ من اليسار إلى اليمين ، إلا إذا اختلفت العلامات وفى هذه الحالة تخضع العمليات لترتيب معين أو أولويات كالتالى :

الرفع للقوة أولاً

يليه الضرب والقسمة وكلاهما فى نفس المرتبة ،

يلى ذلك الجمع والطرح وكلاهما فى نفس المرتبة .

كما يمكن استخدام الأقواس وهذه عند استخدامها تأخذ الأولوية القصوى. تأمل هذه العبارة :

```
30 A = 8 + 4 * 2
```

ناتج هذه العبارة هو 16 وليس24 ، إذ أننا ننفذ، أولاً عملية الضرب 2 * 4 إذ للضرب أولوية على الجمع ، ثم بعد ذلك نضيف 8 للناتج .

أما إذا أردنا أن نضرب 2 فى مجموع 4 + 8 فعلينا فى هذه الحالة أن نضع 4 + 8 بين قوسين وحينئذ تأخذ العملية داخل القوسين الأولوية على العملية خارجها، ويصبح الناتج 24 .

يمكن استخدام تعليمة PRINT التى تطرقنا لها سابقاً لإجراء العمليات الحسابية، وفى هذه الحالة ناتج العملية يتم إظهاره فقط ولا يخزن فى أى متغير.

أمثلة :

```
40 PRINT   6 + 8
50 PRINT   T * 3 - F
60 PRINT   N/ (N - 1)
```

مثال (٢ ، ٢) :

دعنا الآن نكتب برنامجاً كاملاً نستخدم فيه ما تعلمناه حتى الآن . البرنامج التالى يأخذ قيمتين رقميتين يجمعهما، يحسب متوسطهما، ثم يطبع الرقمين والمجموع والمتوسط :

```
10  REM PROGRAM TO ADD TWO NUMBERS
20  REM AND TO PRINT THEIR SUM AND AVERAGE
30  X = 52
40  Y = 34
50  S = X + Y
60  V = S/2
70  PRINT X,Y
80  PRINT S,V
90  END
```

عند تنفيذ هذا البرنامج سيظهر المخرجات التالية :

52	34
86	43

ولا بد أنك لاحظت أن لدينا سطرين فى المخرجات، هذا ناتج عن وجود تعليمتى PRINT فكل تعليمة من تعليمات PRINT تطبع سطراً بعدها. كذلك لا بد أنك لاحظت أن بكل سطر من المخرجات قيمتين، وهذا كذلك ناتج عن وجود عنصرين فى كل من تعليمتى PRINT وأننا قد فصلنا بينهما بواسطة الفاصلة (,).

استخدام الفاصلة (,) كفاصل بين عناصر PRINT يظهر المخرجات وهى بعيدة عن بعضها. فالفاصلة تقسم الشاشة إلى خمس مناطق طباعة، الفرق بين كل منطقة والمنطقة التالية لها يتراوح بين 15 - 19 اعتماداً على الجهاز.

يمكن كذلك استخدام الفاصلة المنقوطة (;) وفى هذه الحالة نجد أن المخرجات تكون قريبة من بعضها.

من الممكن جمع عبارات ثابتة ومتغيرات فى تعليمة PRINT واحدة. فمثلاً زيادة فى الإيضاح وتعريف المخرجات فى البرنامج السابق، كان يمكن للسطر رقم 80 أن يكون كالآتى :

```
80 PRINT 'TOTAL = '; S ;'AVERAGE =' ; V
```

وعندها كانت المخرجات ستكون

```
TOTAL  =  86  AVEARAGE  =  43
```

٤ - أوامر الإدخال :

لإمداد البرنامج بالبيانات فإننا نستخدم تعليمتين - غير تعليمة LET - هما INPUT و READ.

تعليمة INPUT :

نستخدم تعليمة INPUT لإمداد البرنامج بالبيانات أثناء التنفيذ، لذلك فهى مناسبة للبيانات القليلة، وللبرامج التى تتطلب تخاطباً مباشراً بين المستخدم والحاسب.

تتركب التعليمة من عبارة INPUT ثم اسم المتغير، أو أسماء المتغيرات، التى قد تكون متغيرات رقمية أو حرفية.

أمثلة :

```
20 INPUT T
50 INPUT A, B, C
60 INPUT   N, D$, F$, M
70 INPUT   Y$, X
```

تستخدم تعليمة INPUT عندما يراد من المستخدم إدخال بيانات معينة أثناء تنفيذ البرنامج . وعند تنفيذ أى من عبارات INPUT فإنه تظهر على الشاشة علامة استفهام (؟) وما على المستخدم حينئذ إلا إدخال القيمة المطلوبة .

مثال (٢,٣)

```
10 INPUT A
20 PRINT
30 END
```

وعند التنفيذ

```
.RUN
 ? 8
   8
```

وزيادة فى توضيح الشىء المطلوب ، فإنه تستخدم فى العادة تعليمة PRINT قبل أى تعليمة INPUT لتظهر للمستخدم الشىء المطلوب إدخاله .

مثال (٢,٤) :

```
10 PRINT 'ENTER YOUR NAME'
20 NPUT    N$
30 PRINT 'ENTER YOUR AGE'
40 INPUT A
50 PRINT N$; 'YOU ARE' A; 'YEARS OLD'
60 END
```

عند تنفيذ هذا البرنامج :

```
RUN
ENTER YOUR NAME
? ALI
ENTER YOUR AGE
? 23
ALI YOU ARE 23 YEARS OLD
```

تعليمة READ :

تعليمة READ ـ كما يدل على ذلك اسمها ـ تستخدم كذلك لإمداد البرنامج بالبيانات . الفرق بين READ و INPNT أن البيانات فى الأخيرة تدخل للبرنامج أثناء التنفيذ، أما بالنسبة للأولى فإن البيانات تكون مضمنة داخل البرنامج فى عبارة DATA .

مثال : (٥ ، ٢) :

البرنامج التالى يقرأ اسم وعمر شخص، ويطبعهما :

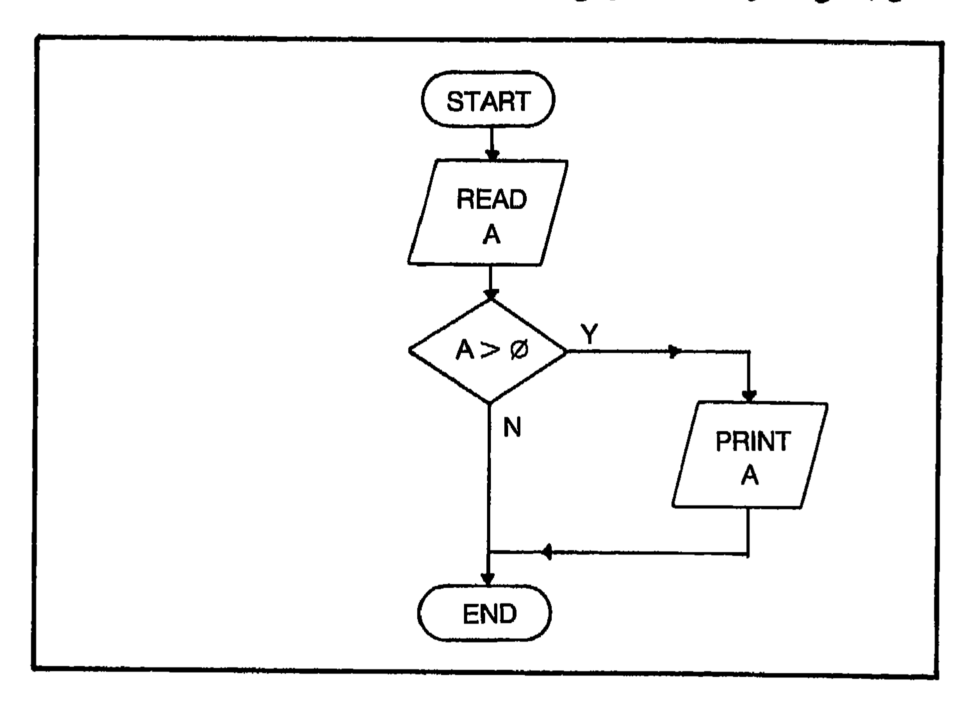
```
10 READ N$, A
20 DATA AHMED, 18
30 PRINT N$; 'IS ' A; 'YEARS OLD '
40 END
RUN
AMED IS 18 YEARS OLD
```

لاحظ أن البيانات موجودة فى عبارة DATA ، كذلك لاحظ الارتباط بين عبارتى READ و DATA . فكل عنصر فى READ لا بد له من بيان فى DATA ويجب أن يكون التسلسل هو نفسه فى العبارتين. عبارة DATA من العبارات التى لا تنفذ، فهى تستخدم فقط لحفظ البيانات؛ لذلك فهى يمكن وضعها فى أى مكان فى البرنامج، ولا يجب بالضرورة أن تكون بعد عبارة READ مباشرة؛ إذ يمكن أن تكون قبلها أو بعدها بعدة عبارات .

٥ ـ نقل التسلسل والمقارنة :

قليل جداً من المسائل هى تلك التى ينساب تسلسلها من أعلى إلى أسفل دون تغيير فى مسارها. معظم المسائل العملية تتطلب اتخاذ قرار معين يغير مسارها فى اتجاه أو آخر،

فمثلاً إذا أردنا أن نقرأ رقماً معيناً، فإن كان الرقم موجباً نطبعه وإلا نهمله، فالمخطط التالى يمكن أن يؤدى ذلك الغرض :



لتمثيل الجـزء الخاص بالمقارنة فى لغة بيسك، فإننا نستخدم أداة المقارنة IF والتى تختبر قيمتين باستخدام الإشارات التالية :

الإشارة	معناها
=	يساوى.
>	أكبر من.
<	أصغر من.
<>	لا يساوى.
>=	أكبر من أو يساوى.
<=	أصغر من أو يساوى.

مثال (٦ , ٢) :

والآن دعنا نكتب البرنامج الذى رسمنا مخططه سابقاً :

```
10  INPUT   A
20  IF A > ∅  THEN 4∅
30  GOTO   5∅
40  PRINT   A
50  END
```

لاحظ أن عبارة IF تتكون من IF ثم الشرط Condition ثم عبارة THEN ثم رقم سطر. فى حالة تحقق الشرط ينفذ الجزء الذى بعد THEN وهو أن ينتقل التسلسل إلى السطر رقم 40 وإذا لم يتحقق الشرط يستمر التسلسل كما هو إلى السطر 30 والذى بدوره ينقل التسلسل بعبارة GOTO إلى السطر الأخير وهو 50 .

تستخدم تعليمة GOTO لنقل التسلسل من سطر إلى سطر آخر، وهو نوعان : نقل غير مشروط، كما فى السطر 30 ، ونقل مشروط أى بشرط معين، كما فى السطر 20 حيث يمكن لتلك العبارة أن تكتب :

```
20 IF A> ∅ GOTO 4∅
```

فعبارتا THEN و GOTO هنا تؤديان نفس الغرض .

من العبارات المطورة عبارة IF ، فليس من الضرورة أن يكون بعد عبارة THEN دائماً رقم سطر، إذ كان يمكن أن نختصر البرنامج السابق كالآتى :

```
10  INPUT A
20  IF A > ∅  THEN PRINT A
30  END
```

وكان سيؤدى لنفس النتيجة ، إلا أن هذا النوع من عبارة IF ليس موجوداً فى كل نسخ بيسك .

٦ ـ الدوارة LOOP وعبارة GOTO :

من الميزات الكبيرة للكمبيوتر قدرته على تكرار عملية معينة ـ أو مجموعة عمليات ـ عدداً من المرات . فمثلاً يمكن تطوير البرنامج فى المثال (٦ , ٢) بحيث يكون لدينا عدد من درجات طلاب فى امتحان ما، ونريد أن نطبع فقط الدرجات من 6ϕ فما فوق. ولنفرض أن نهاية البرنامج تكون بإدخال الدرجة 999 ، فالبرنامج التالى يؤدى ذلك الغرض :

مثال (٢ , ٧) :

```
10   INPUT X
20   IF X    = 999 GOTO 6ϕ
30   IF X   >    = 6ϕ THEN 1ϕ
40   PRINT X
50   GOTO 1ϕ
60   END
```

لاحظ استخدام عبارة GOTO فى السطرين 20 و 50 . عبارة GOTO فى السطر 20 مرتبطة بشرط معين وهو (X = 999) لذلك تعرف GOTO هنا بأنها مشروطة CONDITIONAL ، ويمكن فى مثل هذه الحالات استخدام عبارة THEN كبديل لها، كما هو الحال فى السطر 30 أما GOTO فى السطر 50 فهى غير مشروطة UNCONDITIONAL أى أن مسار البرنامج يتحول من السطر 50 إلى السطر 10 دون شرط. وعلى ذلك فالعبارات من السطر 10 إلى السطر 50 قد كونت لدينا ما يعرف بالدوارة. والدوارة ببساطة هى جزء من البرنامج يتم تنفيذه عدداً من المرات.

البرنامج فى المثال التالى يقوم بقراءة درجات الطلاب، إلا أنه فى هذه المرة يطبع الدرجات أقل من 60 ، كما أن عدد الدرجات يرمز له بالمتغير N والذى يزود به البرنامج أثناء التنفيذ بواسطة عبارة INPUT ليكتسب البرنامج مرونة.

مثال (٢,٨) :

```
10  C   =ɸ
20  INPUT  N  REM
30  INPUT  X        عدد الدرجات
40  IF  X>  = 6ɸ THE N 6ɸ
50  PRINT X
60  C = C + 1
70  IF C < N THEN 3ɸ
80  END
```

لاحظ أننا كونا دوارة بدءاً من السطر 30 وانتهاء للسطر 70 حيث يتم اختيار العداد C فإن كان أقل من عدد الدرجات N فإن التسلسل ينتقل إلى السطر 30 حيث يتم إدخال درجة أخرى.

كذلك لاحظ استخدام العداد C وكيف أننا أعطيناه القيمة صفر فى البداية كإجراء ضرورى لضمان خلوه من أى قيمة سابقة قد تؤثر على النتيجة. ولإجراء عملية العد نفسها، فإننا نضيف واحداً إلى العداد كل مرة بواسطة تعليمة :

```
60   C = C + 1
```

وهى تعنى - كما رأينا فى الفصل السابق - أن قيمة C الجديدة تساوى قيمة C القديمة زائداً واحداً.

تعليمة FOR... NEXT :

يمكن تمثيل الدوارة بصورة أفضل، وذلك باستخدام صيغة FOR...NEXT وهى تغنى عن استخدام العداد واختباره، وما إلى ذلك. الدوارة تتركب من العبارتين التاليتين :

```
20 FOR J  = 1 TO 65 STEP 1
.
.
.
60 NEXT  J
```

لابد لكل عبارة FOR من عبارة NEXT ، والعبارات بين سطرى FOR و NEXT هى التى تكون ما يعرف بجسم الدوارة ويعتمد عدد مرات تنفيذ العبارات فى جسم الدوارة على القيم الموجودة فى عبارة FOR .

تتركب عبارة FOR من متغير رقمى (ل) ثم قيمة ابتدائية (1) ثم قيمة نهائية (65) ثم إضافة (1) ، كما أن عبارة NEXT تتكون من كلمة NEXT ثم متغير، والذى لا بد أن يكون نفس المتغير الذى استخدمناه فى عبارة FOR .

إذن فالتعليمة السابقة تعنى الآتى : نفذ العبارات فى جسم الدوارة من قيمة ل تساوى 1 إلى أن تصبح قيمة ل تساوى 65 مع إضافة 1 إلى ل كل مرة، أى أن الدوارة ستنفذ 65 مرة.

من الجدير ذكره أن الإضافة STEP عندما تكون 1 فإنه يجوز عدم ذكرها، أى أن عبارة FOR السابقة يمكن أن تكون :

```
20    FOR J    = 1 TO 65
```

القيمة الابتدائية، والقيمة النهائية، والإضافة ليس من الضرورى أن تكون أرقاماً ثابتة؛ إذ يمكن أن تكون كلها أو أى منها متغيرات رقمية (معروفاً قيمها مسبقاً). كما أن الإضافة يمكن أن تكون أى قيمة عدد صحيح ، أو كسر عشرى، سالب أو موجب.

أمثلة :

```
10    FOR K = 1 TO N
20    FOR F  = 6 TO 77 STEP A
30    FOR L = J + 5 TO X -  B * D
60    FOR M    = 25 TO 5 STEP - 1
```

مثال (٩,٢) :

البرنامج التالى يقرأ أعمار 50 شخصاً ـ عن طريق الشاشة ـ ويحسب ويطبع متوسط الأعمار :

```
10   T = 0
20   FOR K  = 1 TO 50
30   INPUT A
40   T = T + A
50   NEXT K
60   V  = T/50
70   PRINT V
80   END
```

٧ ـ النسق والمصفوفات :

النسق ARRAY هو عبارة عن مجموعة من البيانات ذات صفة معينة . والفرق بين النسق والمتغير العادى ، أن كل عنصر من عناصر البيانات فى النسق تكون له خانته المحجوزة فى الذاكرة وبالتالى يمكن الرجوع لأى من هذه العناصر فى أى جزء من البرنامج ، دون إعادة قراءة البيانات .

يتم التعبير عن النسق بعبارة DIM . هذه العبارة تؤدى لحجز خانات فى الذاكرة بالعدد المذكور فى عبارة DIM .

فمثلاً العبارة :

```
30   DIM R (35)
```

تؤدى إلى حجز 35 خانة فى الذاكرة لمتغير رقمى اسمه R . وفى هذه الحالة ، يتم الرجوع لأى عنصر من عناصر R بواسطة ترتيبه داخل النسق ، فأول عنصر يأخذ الاسم (1) R والعنصر التاسع يأخذ الاسم (9) R والعنصر الثلاثون يأخذ الاسم (30) R وهكذا . ولا بد أن يكون الرقم بين قوسين ، هذا الرقم يعرف بالمؤشر SUBSCRIPT ، ولا بد للمؤشر أن يكون بين 1 والحد الأقصى للنسق والمذكور فى عبارة DIM ـ فى هذا المثال ـ 35 .

مثال (٢,١٠) :

البرنامج التالى يقرأ أطوال 20 لاعباً فى فريق لكرة السلة ويقوم بطباعة أكبر طول .

```
10   DIM  L REM(20)      لحجز الخانات
20   H = 0 REM     لحفظ أكبر طول
30   FOR P = 1 TO 20
40   INPUT  L ( P)
50   NEXT P
60   FOR V = 1 TO 20
70   IF L (V) H THEN  H = L (V)
80   NEXT V
90   PRINT H REM          الطول الأكبر
100  END
```

النسق ذو البعدين :

حديثنا عن النسق حتى الآن كان عما يعرف بالنسق ذى البعد الواحد ، إلا أن النسق يمكن أن يكون له أكثر من بعد. إذ يمكن أن يكون لدينا نسق ذو بعدين ، وهذا ما يعرف كذلك بالمصفوفة ، والتى يتم التعبير عنها كذلك بعبارة DIM ولكن فى هذه الحالة يذكر رقمان ، الأول يمثل عدد الصفوف ، والثانى عدد الأعمدة.

فمثلاً متوسط درجات الحرارة لثلاث مدن خلال 12 شهراً يمكن أن تمثل مصفوفة ذات 3 صفوف و 12 عموداً ، حيث يمثل كل صف إحدى المدن الثلاث ، وكل عمود يمثل أحد الشهور الاثنى عشر. للتعبير عن هذه المصفوفة نستخدم تعليمة DIM كالآتى :

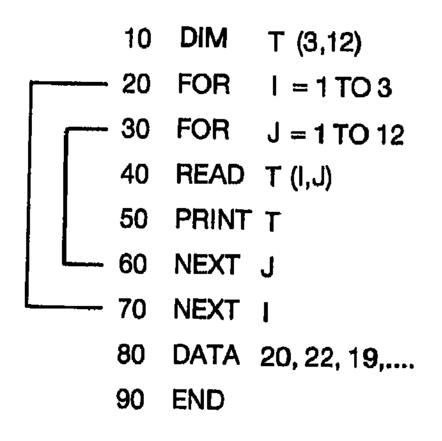
```
10   DIM     T (3, 12)
```

هذه العبارة تؤدى إلى حجز 36 خانة عبارة عن 3 صفوف و 12 عموداً لمصفوفة ذات بعدين باسم T .

للرجوع لأى من عناصر هذه المصفوفة ، لا بد من ذكر اسم المصفوفة ، ثم مؤشرين لتحديد العنصر المطلوب. يحدد العنصر الأول رقم الصف ، ويحدد الثانى رقم العمود.

دعنا نرى الآن كيف تعامل هذه المصفوفة ، فمثلاً إذا أردنا قراءة المصفوفة وطباعتها فإننا نكتب البرنامج التالى :

مثال (١١، ٢) :



```
10  DIM    T (3,12)
20  FOR    I = 1 TO 3
30  FOR    J = 1 TO 12
40  READ   T (I,J)
50  PRINT  T
60  NEXT   J
70  NEXT   I
80  DATA   20, 22, 19,....
90  END
```

لعلك لاحظت أن لدينا فى البرنامج عبارتى FOR وعبارتى NEXT وهذا ما يعرف بالدوارات المتداخلة NESTED LOOPS . ولكى نستطيع التعامل مع كل عناصر النسق ذى البعدين لا بد من وجود هذا النوع من الدوارات.

عند وجود دوارتين متداخلتين، فالقاعدة هى أن الدوارة التى تبدأ أولاً تنتهى أخيراً، والتى تبدأ أخيراً تنتهى أولاً. فدوارة I فى المثال بدأت أولاً، لذلك انتهت أخيراً، فى حين أن دوارة J بدأت أخيراً وانتهت أولاً. هذه القاعدة ضرورية لضمان عدم تقاطع دوارتين.

أوامر المصفوفات :

نسبة لما للمصفوفات من كثرة استخدام، فقد أفردت لها أوامر خاصة بها فى لغة بيسك باستخدام عبارة MAT .

لقراءة مصفوفة ما، فإننا نذكر عبارة MAT ثم اسم المصفوفة فقط، فمثلاً المصفوفة التى بالمثال السابق يمكن قراءتها بالعبارة التالية :

```
20   MAT   READ   T
```

هذه العبارة تؤدى لقراءة جميع عناصر المصفوفة T .

ولطباعة المصفوفة نستخدم هذه التعليمة :

```
30   MAT   PRINT   T
```

كذلك يمكن إجراء العمليات الحسابية على المصفوفات، فلجمع مصفوفتين B, A فى مصفوفة ثالثة C فإننا نستخدم هذه العبارة :

```
30   MAT   C  = A + B
```

نفس الشىء إذا أردنا طرح مصفوفتين ووضع الناتج فى ثالثة :

```
40   MAT   C  = A - B
```

هنالك قاعدة هامة لجمع أى مصفوفتين أو طرحهما، هى أنه فى كلتا الحالتين لابد من أن تكون المصفوفات الثلاث لها نفس الأبعاد (أى نفس عدد الصفوف، ونفس عدد الأعمدة) يعنى أنه فى العبارتين السابقتين لابد أن تكون المصفوفات C, B, A لها نفس الأبعاد.

ضرب المصفوفات :

يمكن كذلك ضرب مصفوفتين P, Q ووضع الناتج فى مصفوفة R بالعبارة التالية :

```
30   MAT   R  = P*Q
```

كما للجمع والطرح قاعدة، فللضرب كذلك قاعدة، هى أنه عند ضرب مصفوفتين لابد أن يكون عدد أعمدة المصفوفة الأولى (P) مساوياً لعدد صفوف المصفوفة الثانية (Q)، أما المصفوفة الناتجة فلا بد أن يكون لها نفس عدد صفوف المصفوفة الأولى (P) ونفس عدد أعمدة المصفوفة الثانية (Q).

مدور المصفوفة TRANSPOSE :

مدور المصفوفة هو مصفوفة أخرى، صفوفها هى أعمدة المصفوفة الأولى، وأعمدتها هى صفوف المصفوفة الأولى. للحصول على مدور المصفوفة فإننا نستخدم عبارة TRN .

مثال :

```
60   MAT   H  = TRN (F)
```

وعند كتابة عبارات DIM فى هذه الحالة لابد أن تكون أبعاد المصفوفة H عكس أبعاد المصفوفة F مثلاً :

```
20   DIM   F  (3,4), H (4,3)
```

مقلوب المصفوفة INVERSE :

مقلوب المصفوفة هو مصفوفة أخرى، تحمل نفس الأبعاد، وللحصول على مقلوب المصفوفة، لابد أن تكون المصفوفة مربعة SQUARE . (ليس من الضرورى أن يكون لكل مصفوفة مربعة مقلوب).

مثال :

إذا كانت A مصفوفة ذات أبعاد 3 × 3 فإن عبارة :

```
30   MAT   B  = INV (A)
```

ستحسب مقلوب المصفوفة A وتضعه فى المصفوفة B (وهى مصفوفة أخرى ذات أبعاد 3 × 3).

هنـالك بالطبع تعليمات أخرى خاصة بالمصفوفات، وهذه تختلف من جهاز لآخر، وما تعرضنا له هنا هو أهم التعليمات.

٨ ـ الدوال FUNCTIONS:

هنالك العديد من العمليات التى يكثر استخدامها فى البرامج، وقد يتطلب حسابها أو إجراؤها كتابة عدد من التعليمات، لذلك فقد تم تخزين هذه العمليات ضمن مترجم اللغة، وذلك لتسهيل كتابة البرامج. هذه التعليمات تعرف بالدوال، ويتم استخدامها فى البرنامج عن طريق استدعائها.

لاستـدعـاء دالـة معينـة فى البرنامج فإنك تكتب اسمها (والذى يتكون عادة من ثلاثة أحرف)، ثم التعبير الذى تريد إجراء العملية عليه بين قوسين. من أكثر الدوال استخداماً دالة الجذر التربيعى، واسمها SQR ، فمثلاً إذا أردنا أن نحسب الجذر التربيعى للرقم 64 وطباعته فإننا نكتب :

```
10  D  =SQR (64)
20  PRINT  D
```

أو مباشرة :

```
10  PRINT  SQR  (64)
```

ويمكن للتعبير أن يكون متغيراً مثل :

```
30  V  =SQR (P)
```

أو تعبيراً حسابياً مثل :

```
10  PRINT  SQR (T - N + 1)
```

هنـالـك العديد من هذه الدوال تختلف من جهاز لآخر، بعضها دوال رقمية، وبعضها حرفية، وبعضها خاص بحساب المثلثات، مثل : الظل والجيب . . . إلخ. ولمعرفة الدوال الموجودة فى أى جهاز ينبغى الرجوع للوثائق الخاصة بذلك الجهاز (ملحق رقم ١٢).

٩ ـ دوال المبرمج :

بإمكان المبرمج أن يعرف دوال خاصة به، غير تلك التى توفرها اللغة.

تعريف الدالة :

لتعريف أى دالة فإنك تكتب عبارة DEF ثم اسم الدالة، والذى يتركب من ثلاثة أحرف، الحرفان الأولان منها هما FN ثم أى حرف آخر.

أمثلة :

FNA

FND

FNX

بعد ذلك تعبير الدالة بين قوسين ثم علامة = ثم التعبير المطلوب إسناده كمحصلة للدالة.

```
20   DEF   FNA( X) = X ** 3
```

هذه الدالة تكون محصلتها رفع قيمة المتغير X إلى القوة 3 ، علماً بأن المتغير X هنا يدل على أى متغير رقمى، ولا علاقة لاسمه بالاسم الذى سوف تستدعى به الدالة كما سنرى.

استدعاء الدالة :

لاستدعاء الدالة فإنك تذكر اسمها، ثم التعبير بين قوسين، وعندها يتم التعويض عن X فى الدالة بالمتغير.

مثال :

إذا كان لدينا الدالة السابقة فمن الممكن استدعاؤها كالآتى :

```
50   PRINT   FNA ( 5)
```

وعندها سيتم طباعة الرقم 125 وهو عبارة عن 5 مرفوعة للقوة 3 .

أو يمكن أن نكتب الآتى :

```
40   N  = 3
50   C  = FNA ( N)
```

أما هنا فإن المتغير C سيأخذ قيمة المتغير N مرفوعاً للقوة 3 وهى 27 .

مثال :

```
10 DEF FNK ( S) = S ** ( 1/3)
20 P = 216
30 PRINT FNK ( P)
40 END
```

العبارة بالسطر 10 دالة تقوم بحساب الجذر التكعيبى . السطر 30 يستدعى الدالة بتعويض المتغير P بدلاً عن S وحيث إن قيمة P تساوى 216 فسيتم حساب وطباعة الجذر التكعيبى للرقم 216 وهو 6 .

يمكن للدالة أن تأخذ أكثر من متغير واحد .

مثال :

```
10 DEF FNR ( J, K, L) = ( J + K + L) /3
```

هذه الدالة محصلتها متوسط ثلاث قيم ، ويمكن استخدامها لاستخراج المتوسط لأى ثلاثة أرقام كالآتى :

```
10 DEF FNR ( J, K, L) = ( J + K + L) /3
20 A = 7
30 B = 2
40 C = 6
50 D = FNR ( A, B, C)
```

المتغير D فى السطر 50 سيأخذ القيمة 5 وهى محصلة الدالة FNR وهى جمع القيم الثلاث وقسمتها على 3 كما فى تعريف الدالة فى السطر 10 .

١٠ ـ عبارات إخراج متقدمة :

كما رأينا فإن استخدام الفاصلة كفاصل بين العناصر فى عبارة PRINT يجعل المخرجات مفرقة فى مناطق محددة ، كما أن الفاصلة المنقوطة تضم العناصر بعضها مع بعض . قد يحتاج المبرمج أحياناً إلى أن تكون المخرجات بشكل يحدده هو بمرونة أكثر مما هو موجود باستخدام الفاصلة والفاصلة المنقوطة . لتحقيق هذا الهدف هنالك طريقتان هما :

```
PRINT TAB
PRINT USING
```

عبارة PRINT TAB :

باستخدام TAB فى أمر PRINT يمكن للمبرمج أن يحدد فى أى عمود من السطر يريد أن يظهر كل عنصر من عناصره .

مثال :

```
20  N $ = 'MEAN = '
30  PRINT   TAB ( 5) ; N$
40  END
```

عند تنفيذ هذا البرنامج ستظهر عبارة = MEAN عند العمود رقم 5 .

عبارة PRINT TAB تتكون من العبارة نفسها ، ثم رقم العمود الذى يراد الطباعة عنده ، ثم فاصلة أو فاصلة منقوطة ، ثم العنصر المراد إظهاره ، متغيراً كان أم ثابتاً ، رقمياً كان أو حرفياً .

أمثلة :

```
10  PRINT   TAB ( 6) , A
20  PRINT   TAB ( 12) ; B
30  PRINT   TAB ( 2) ; 'PROGRAM 5'
40  PRINT   TAB ( 30) , 'TOTAL IS'
```

كما يمكن جمع أكثر من عنصر فى عبارة PRINT TAB واحدة ، ونحدد لكل عنصر العمود الذى يطبع بدءاً منه .

مثال :

```
10  A$  = 'FIELD A IS'
20  B$  = 'FIELD B IS'
30  A   = 16
40  B   = 4
50  PRINT   TAB ( 3) , A$; TAB ( 15) ; A; TAB (2Ø), B$ ; TAB ( 32) ;B
60  END
```

وعند تنفيذ هذا البرنامج فإنه سيخرج المخرجات التالية :

عمود 3	عمود 15	عمود 20	عمود 32

```
FIELD  A  IS    16    FIELD  B  IS        4
```

ولعلك لاحظت أن الأرقام تسبق بفراغ وذلك للإشارة (+ أو −).

عبارة PRINT USING :

تمكن عبارة PRINT USING المبرمج من التحكم أكثر فى شكل المخرجات، وتتركب العبارة من سطرين : الأول يحتوى على عبارة PRINT USING مع رقم السطر الذى به شكل الطباعة المطلوبة، والثانى يمثل شكل الطباعة المطلوبة، والذى يبدأ بعلامة (:)

مثال :

```
10  A  = 64
20  PRINT   USING 3Ø, A
30  : # #
40  END
```

المقصود هنا فى السطر 2Ø هو أمر بطباعة قيمة المتغير A باستخدام الشكل الذى فى السطر رقم 3Ø . والسطر رقم 3Ø يحتوى على علامة : ثم علامتى ## هذه العلامات تمثل كل واحدة منها خانة من المتغير المراد طباعته، رقمياً كان أم حرفياً.[1]

عند تنفيذ هذا البرنامج فإنه سيخرج المخرجات التالية :

```
64
```

مثال :

```
10  A = 12
20  PRINT   USING 30, A, A*A
30  : THE SQUARE OF # # IS # # #
40  END
```

(١) ينبغى الرجوع للوثائق الخاصة بأى جهاز لمعرفة العلامات المستخدمة فى هذه العبارة.

عند تنفيذ هذا البرنامج سيخرج المخرجات التالية :

```
THE   SQUARE   OF   12  IS   144
```

لاحظ كيف أننا جمعنا بين الثوابت والمتغيرات فى السطر 30 ولاحظ كذلك أنه عند كتابة أى ثوابت فى سطر الشكل، فإنها لاتكون بين علامتى ' ' .

كذلك يمكن طباعة متغيرات حرفية بنفس عبارة USING :

مثال :

```
10  N$ = 'JANUARY'
20  PRINT USING 3∅, N$
30  : THE FIRST MONTH OF THE YEAR IS # # # # # # #
40  END
```

ومخرجات هذا البرنامج ستكون :

```
THE FIRST MONTH OF THE YEAR IS JANUARY
```

إذا زادت علامات # عن حجم المتغير المراد طباعته، فلا توجد مشكلة، أما إذا نقصت فستستبدل القيمة المراد طباعتها بعلامة * .

مثال :

```
10  A = 12
20  PRINT USING 3∅, A , A*A
30  : THE SQUARE OF   ##   IS   ##
40  END
RUN
THE   SQUARE   OF   12   IS **
```

عبارة PRINT USING يمكن ألا تشتمل على متغيرات أو عناصر، فقط رقم سطر يحتوى على قيمة ثابتة (عنوان مثلاً) .

مثال :

```
30   PRINT USING 6 0
.
.
60   :      EMPLOYEE REPORT
```

عند تنفيذ السطر 30 فإنه سيتم طباعة القيمة الموجودة فى السطر 60 وهى :

EMPLOYEE REPORT

ولعلك لاحظت أن سطر الشكل لا يتبع بالضرورة السطر الذى يستدعيه (وهذا منطقى ، لأن نفس السطر يمكن أن يستدعى من عدة مناطق فى البرنامج).

يمكن أن تحتوى القيم الرقمية على خانات عشرية ، وفى هذه الحالة فإن الرقم فى سطر الشكل يتم توصيفه مع وضع الفاصلة العشرية (.) فى المكان المناسب.

مثال :

```
10   P = 55
20   PRINT USING 30, P, SQR ( P)
30   : THE SQUARE ROOT OF ##IS #.###
40   END
```

ومخرجات هذا البرنامج ستكون :'

THE SQUARE 'ROOT OF 55 15 7.412

١١ ـ البرامج الفرعية SUBROUTINES :

البرنامج الفرعى هو جزء من برنامج بيسك ، يتم تعريفه فى مكان ما من البرنامج ، ويمكن استدعاؤه من عدة أماكن فى البرنامج.

البرنامج قد يحتوى على أكثر من برنامج فرعى ، والبرنامج الفرعى لايبدأ بأى عبارة معينة ، إلا أنه ينتهى بعبارة RETURN والتى لها الشكل العام التالى :

رقم السطر RETURN

وبما أننا ذكرنا أن البرنامج الفرعى يمكن استدعاؤه من عدة أماكن فى البرنامج ، فعبارة RETURN تعيد التسلسل إلى المكان الذى تم فيه استدعاء البرنامج الفرعى.

استدعاء البرنامج الفرعى يكون بعبارة GOSUB وبعدها يذكر رقم السطر الذى يبدأ فيه البرنامج الفرعى . الشكل العام للتعليمة هو :

رقم سطر GOSUB رقم سطر

فيما يلى هيكل لبرنامج بيسك يستخدم برنامجاً فرعياً :

```
10 ..
20 ..
:
40      GOSUB 70
50 ..
60      STOP
70      REM SUBROUTINE FOR ....
80
90
100      RETURN
110 ..
.
.
.
```

بداية البرنامج الفرعى هنا عند السطر 70 والعبارات من 70 إلى 90 تمثل العبارات الخاصة بالبرنامج الفرعى . عبارة RETURN عند السطر 100 تمثل نهاية البرنامج الفرعى .
يتم استدعاء البرنامج الفرعى عند السطر 40 بعبارة GOSUB .

لعلك لاحظت وجود عبارة STOP عند السطر 60 هذه تمثل النهاية المنطقية للبرنامج ، وليس من الضرورة أن ينتهى البرنامج بعبارة END (إذ أن هذه العبارة ـ كما قدمنا ـ تدل على النهاية المادية للبرنامج) .

فى العادة يكتب البرنامج الفرعى أو البرامج الفرعية فى آخر البرنامج ، ويتم استدعاؤها فى الجزء العلوى للبرنامج ، وفى هذه الحالة لا بد من وجود عبارة STOP أو أمر GOTO إلى سطر به عبارة END فى نهاية البرنامج .

مثال :

الـبرنـامج التالى مكتوب بطريقة البرامج الفرعية، وهو يقوم بقراءة أسماء عشرة طلاب، ودرجات كل منهم فى 3 اختبارات، ومن ثم يحسب المجموع والمعدل لكل طالب ويطبع كل البيانات.

```
10  REM  برنامج لقراءه اسماء ودرجات 10 طلاب في 3 اختبارات
20  REM  وحساب مجموع ومعدل كل منهم وطباعتها. البرنامج
25  REM  SUBROUTINE يستخدم طريقه البرامج الفرعيه
30  DIM N$(10),A(10),B(10),C(10),T(10),V(10)
40  PRINT USING 400
50  PRINT USING 410
55  PRINT
60  GOSUB 100
70  GOSUB 200
80  GOSUB 300
90  STOP
100 FOR I=1 TO 10
130  READ N$(I),A(I),B(I),C(I)      REM  الاسم والدرجات الثلاث
140 NEXT I
150 RETURN
200 FOR I=1 TO 10
210  T(I)=A(I)+B(I)+C(I)  REM  المجموع
220  V(I)=T(I)/3   REM  المعدل
230 NEXT I
240 RETURN
300 FOR I=1 TO 10
310  PRINT USING 500,V(I),T(I),C(I),B(I),A(I),N$(I),I
320  PRINT USING 510
330 NEXT I
340 RETURN
400 :   معدل     مجموع     درجه٣    درجه٢    درجه١     الاســـــم          م
410 :   ________________________________________________________________
500 :   ##.##     ###       ###      ###      ###   ################  (##
510 :   ________________________________________________________________
600 DATA حسن صالح,82,91,88,ملاح القاضي,75,72,76,عبد العزيز علي,80,82,85
610 DATA سالم فارس,81,75,88,احمد العبيسي,65,52,60,محمد الحميدان,90,88,86
620 DATA كمال ابوراس,89,91,93,عمر حسن,75,76,80,جمال عثمان,85,84,88
630 DATA طلال الناجي,89,91,93
999 END
```

المخرجات

م	الاســـــم	درجه١	درجه٢	درجه٣	مجموع	معدل
(1	حسن صالح	82	91	88	261	87.00
(2	ملاح القاضي	75	72	76	223	74.33
(3	عبد العزيز علي	80	82	85	247	82.33
(4	سالم فارس	81	75	88	244	81.33
(5	احمد العبيسي	65	52	60	177	59.00
(6	محمد الحميدان	90	88	86	264	88.00
(7	كمال ابوراس	89	91	93	273	91.00
(8	عمر حسن	75	76	80	231	77.00
(9	جمال عثمان	85	84	88	257	85.67
(10	طلال الناجي	89	91	93	273	91..00

تمارين

١) القـائمـة التـالية تحتـوى على متغـيرات رقمية وحرفية، بعضها خطأ . المطلوب تحديد المتغيرات الصحيحة ونوعها - (رقمى، حرفى) - والمتغيرات الخطأ والسبب :

1 – Y3	2 – E$	3 – $#	4 – R + 9
5 – DD	6 – R 8	7 – 'B5'	8 –X
9 – SA	10 – NM		

٢) اكتب عبارات LET للمعادلات الجبرية التالية :

1 – $J = (a/b) + 3$

2 – $L = L/(M + 8)$

3 – $C = (P + Q)(R - V)^2$

4 – $X = 5Y - 2$

5 – $Z = X^2 + Y^2$

٣) اكتب عبارات PRINT للآتى :

١ - اطبع العبارة CORRECT ANSNER فى العمود الأول.

٢ - اطبع الرقم 5 فى المنطقة الأولى، والرقم 6 فى المنطقة الثانية.

٣ - اطبع السطرين التاليين باعتبار أن 16 و 36 هما قيمتا المتغيرين A و B على التوالى :

```
THE   FIRST    ANSWER  IS  16
THE   SECOND   ANSWER  IS  36
```

٤) اكتب العبارتين فى السؤال (٣) ولكن فى سطر واحد كالآتى :

```
ANSWER  1  IS 16  WHILE  ANSWER  2  IS  36
```

٥) بافتراض أن قيمة X هى 3 وقيمة Y هى 15 اطبع ناتج ضرب الرقمين كالآتى :

```
THE   MULTIPLE   OF  3  AND  15  IS  45
```

٦) أوجد الخطأ فى الأجزاء التالية من البرامج :

```
1 -  50  INPUT  XY
2 -  20  INPUT  A  B  C
3 -  10  INPUT  C, D, E
      ?  5  9  33
4 -  60  INPUT  J, K
      ?  84
5 -  80  INPUT  M, N$
      ?  ALI, 36
6 -  40  READ   A, B, C
     50  DATA   169,32
7 -  20  DATA   135, ZAID
     30  READ   A$, B
8 -  50  DATA   39, 50
     60  READ   J, K, L
9 -      DATA   6, 3
     20  READ   X, Y
10 - 10  READ   8
     20  DATA   7
```

٧) اكتب كلاً من البرامج التالية بطريقتين : الأولى باستخدام أمر INPUT والثانية باستخدام READ.

١ ـ اكتب برناجاً لقراءة رقمين، ثم طباعة مجموعهما.
٢ ـ اكتب برناجاً لقراءة ثلاثة أرقام، وطباعة مجموعها ومتوسطها.
٣ ـ اكتب برناجاً لقراءة رقم يمثل المسافة وآخر يمثل الزمن ومن ثم طباعة السرعة.
٤ ـ اكتب برناجاً لتحويل أى عدد من الأميال إلى ما يقابله بالكيلومترات، علماً بأن الميل = 9/5 كيلو متر.
٥ ـ اكتب برناجاً لقراءة ثلاثة أرقام، وطباعتها، وطباعة مربعها.

٨) أوجد الخطأ فى العبارات التالية :

```
1- 60   IF   N$ = NO    THEN 30
2- 40   IF   A + J = 6  GOTO 70
3- 30   IF   K = L + 1  THEN  GOTO  10
4- 40   GOTO  20  IF  C = 9
5- 10   FOR   I = 5 TO 1
6- 20   FOR   J = 1 TO 20
   30   K  = K + J
   40   PRINT  K
   50   END
7- 30   FOR  F $ = 1 TO 7
8- 10   FOR  K = 1 TO 5
   20   INPUT X
   30   IF X>99 THEN 10
   40   NEXT K
9- 70   NEXT 30
10-   FOR   J = 5 TO 10 STEP - 1
```

٩) اكتب برنامجاً لقراءة رقمين J,K ثم طباعة الرقم الأكبر أولاً ، ثم الرقم الأصغر.

١٠) اكتب برنامجاً لقراءة 10 أرقام وطباعة الأرقام الموجبة فقط .

١١) اكتب برنامجاً لقراءة 25 رقماً تمثل أعمار طلاب فى أحد الفصول ، ثم طباعة العمر الأكبر.

١٢) اكتب البرنامج فى السؤال (١١) ، ولكن طباعة العمر الأصغر.

١٣) أوجد الخطأ فى العبارات التالية :

```
1- 30   DIM   X,Y, Z (25)
2- 50   DIM   N1 (15)
3- 90   DIM   K
4- 30   DIM   F (I,J)
5- 10   DIM   T (3,2)
   20   READ   T (2)
```

١٤) اكتب عبارة DIM لحجز 15 خانة لنسق يمثل عدد سكان 15 مدينة .

١٥) اكتب عبارة DIM لحجز نسق يمثل أسماء 20 عاصمة .

١٦) اكتب عبارةDIM لحجز نسق ذى بعدين، يمثل مبيعات4 أفرع لإحدى الشركات خلال 12 شهراً.

١٧) اكتب برنامجاً لقراءة بيانات تمثل درجات 20 طالباً فى 4 اختبارات، ثم حساب وطباعة المعدل لكل طالب، وطباعة المعدل العام.

١٨) اكتب برنامجا لقراءة مصفوفةX وهى 4 × 3 ثم طباعة مجموع كل صف (لاتستخدم عباراتMAT).

١٩) اكتب البرنامج ١٨ ولكن مع طباعة كل عمود.

٢٠) اكتب المطلوب فى البرنامجين (١٨) , (١٩) معاً فى برنامج واحد، وباستخدام عباراتMAT .

٢١) البرامج التالية تحتوى على أخطاء، حددها :

```
1 -   10  INPUT X,Y
      20  IF   X> Y THEN 50
      30  PRINT *** 
      40  STOP
      50  PRINT Y*Y
      60  RETURN
      70  END


2 -   10  INPUT K, L
      20  IF K + L > 5 GOSUB 70
      30  PRINT K, L
      40  GOTO 10
      50  END


3 -   10  INPUT M
      20  GOSUB
      30  PRINT M + 1
      40  GOTO 10
      50  PRINT M*M
      60  END
```

٢٢) اكتب برنامجاً بطريقة البرامج الفرعية للآتى :
إحدى المحلات الكبرى لديها برنامج خصم لزبائنها حسب قيمة مشترياتهم كالآتى :

قيمة المشتريات	نسبة الخصم
أقل من 100	لا يوجد
100 – 299	3%
300 – 499	5%
500 – 799	8%
800 فأكثر	12%

المطلوب قراءة البيانات التالية لكل زبون :
- رقم الزبون .
- إجمالى قيمة المبيعات .

ومن ثم حساب الخصم لكل زبون، وطباعة تقرير يشتمل على :
- رقم الزبون .
- إجمالى قيمة المبيعات .
- قيمة الخصم .
- صافى القيمة المدفوعة .

وفى نهاية البرنامج اطبع :
- إجمالى المبيعات .
- إجمالى الخصم .
- إجمالى الصافى .

اكتب البيانات لـ 20 زبوناً .

٢٣)تعرف على الأخطاء فى الأجزاء التالية من البرامج :

1 -
```
10 A = 6
20 B = 8
30 PRINT X, TAB (9), Y
40 END
```

2 -
```
10 N $ = 'WRONG'
20 PRINT ANSWER FOR Q. Z IS' TAB (6); N$
30 END
```

3 -
```
10 A = 65
20 B = 94
30 C = 283
40 PRINT USING 50, A, B, C
50 :  ##   # # #
60 END
```

4 -
```
10 X = 3
20 Y = 5
30 PRINT USING 40, X, Y
40 # #   # #
50 END
```

5 -
```
10 READ J, K
20 L = J* K
30 PRINT USING 50, J, K, L
40 DATA 13, 19
50 : 'MULTIPLE OF'; J;' A N D'; K; 'IS' L
60 END
```

6 -
```
10 READ M, N
20 DATA 66, 55
30 PRINT USING 4O, M, N, M *N
40 :  ##     ##     ###
50 END
```

٢٤)ماذا ستكون مخرجات كل من البرامج التالية :

1 -
```
10 DEF FNC (A) = A**3 + 2 * A**2 - 3*A + 4
20 T = 4
30 PRINT FNC (T)
40 END
```

2 -
```
10 DEF FND (Y) = Y/3 + /2 + Y
20 X = 18
30 PRINT FND (X) + 1
40 END
```

3 -
```
10 DEF FNN (P,Q) = P**2 + Q**2
20 J = 3
30 K = J + 4
40 R = FNN (J,K)
50 PRINT R
60 END
```

4 -
```
10 DEF FNY (S, T, U) = SQR (S**2 + T**2 + U**2)
20 A = 6
30 B = 7
40 C = 8
50 D = FNY (A, B, C)
60 PRINT D
70 END
```

٢٥) اكتب تعاريف دوال (DEF FN) للآتى :

١ ـ المجموع الكلى لرقم، ومربعه ومكعبه.

٢ ـ المتوسط لثلاثة أرقام.

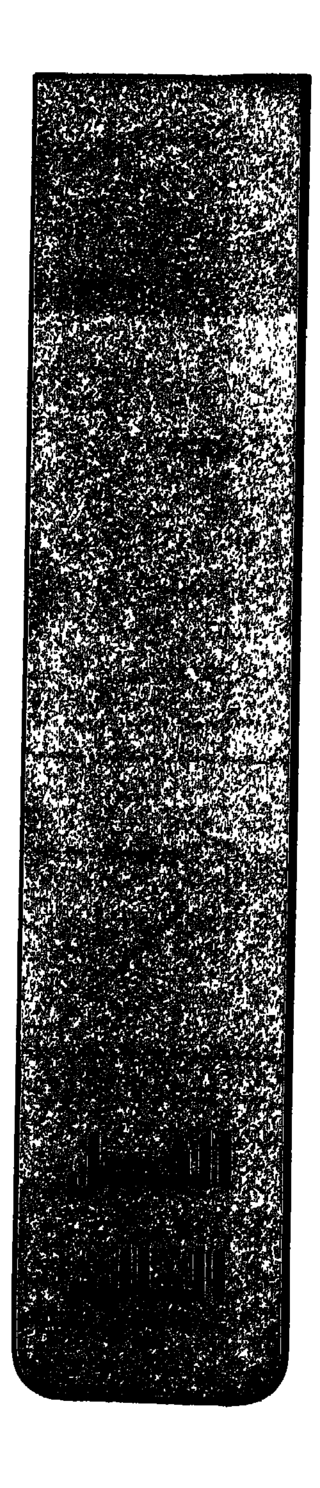
٣ ـ محيط الدائرة بدلالة نصف القطر.

٤ ـ مساحة الدائرة بدلالة نصف القطر.

٥ – قيمة المعادلة

$$T = \frac{(P/Q) + (R/S)}{2}$$

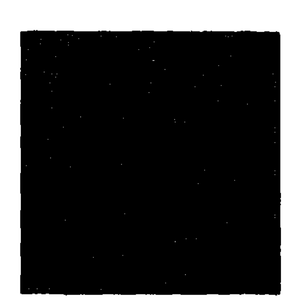




التوزيعات التكرارية لبيانات العينة



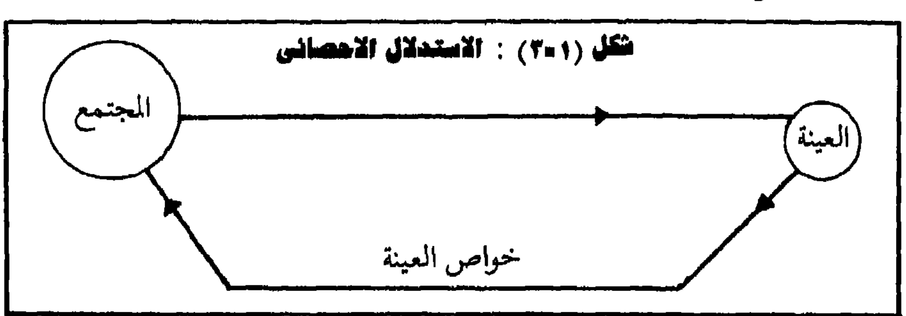




التوزيعات التكرارية لبيانات العينة

١-١ المقدمة :

تسمى جميع نتائج التجربة الإحصائية بمجتمع النتائج (المجتمع) أو البيانات، وأى جزء من هذه البيانات يسمى ببيانات العينة، وبيانات العينة هى المشاهدات (المتغيرات) التى يمكن دراستها وتحليلها لمعرفة خواص المجتمع المعنى بالدراسة. ومن هنا جاء مفهوم الاستدلال الإحصائى (Statistical Inference) الذى يتكون من مجموعة الاستنتاجات الخاصة بالمجتمع الإحصائى، وتعتمد هذه الاستنتاجات فى جملتها على المعلومات التى استخلصت من بيانات العينة.



شكل (١-٢) : الاستدلال الاحصائى

إذاً فالمشكلة الأساسية هى معرفة خواص العينة للوصول لخواص المجتمع الذى سحبت منه العينة.

يختلف نوع ومستوى دقة الخواص المستخرجة باختلاف صفات البيانات، والنموذج الملائم للتطبيق، فقد يستطيع المحلل فى بعض الحالات تحديد ذلك النموذج بناء على معلومات نظرية، أو قياساً على حالات تجريبية مشابهة. ويصبح كل المطلوب فى هذه الحالة هو استخدام البيانات، لتقدير المعالم الخاصة بالعينة والمجتمع - كالوسط الحسابى والانحراف المعيارى اللذين سيرد ذكرهما فيما بعد.

بيد أن المحلل فى كثير من الحالات لا يستطيع تحديد النموذج الملائم مسبقاً، لعدم توفر معلومـات نظرية ، أو حالات تطبيقية تلائم البيانـات التى بين يديه تماماً. إذاً فلا بد من استخـدام هذه البيانات لتحديد الخواص الأساسية لأفضل النماذج كمرحلة أولية فى تحليل البيانات. وهذا هو الهدف من هذا الفصل.

٢ ـ أنواع البيانات :

تنقسم البيانـات إلى نوعـين، هما : البيانـات الكمية (QUANTITATIVE) ، والبيانات الـوصفية أو النوعية (QUALITATIVE) . فالبيانات الوصفية هى التى تصنف البيانات إلى صفات معينة، كالجنسيات، والمناطق، والحالة الاجتماعية، والحالة الوظيفية.

مثال (١ ,٣) :

مثال لبيانات وصفية :

الجدول التالى يوضح النسب المئوية للمشتغلين الذين تبلغ أجورهم الأسبوعية ٧٠٠ ريال فأكثر داخل كل مجموعة من مجموعات المهن فى المؤسسات الخاصة*.

مجموعات المهن	الرياض		جدة		الدمام	
	سعوديون	غير سعوديين	سعوديون	غير سعوديين	سعوديون	غير سعوديين
المهن الفنية والعلمية	٦١,٤	٧١,٧	٦٤,٣	٦٣,٧	٧٥,٤	٥٠,٥
رؤساء وأعضاء المجالس والمديرون	٩٨,٠	٩٥,٥	٩٥,٩	٨٨,٢	٩٤,٤	٨٦,٩
الأعمال الكتابية	٤١,٦	٤٧,٣	٤٥,٦	٤١,٦	٦١,١	٣٩,٢
القائمون بأعمال البيع والشراء	٤٨,٢	٤١,١	٥٦,٥	٥١,٩	٦٤,٠	٤٨,٦
المشتغلون بالخدمات	٤,٢	٣,٧	٥,٩	٢,٥	١,٨	٢,١
وسائل النقل	٣٧,١	١١,٣	١٩,١	٩,٢	٤,٤	٧,١
المجموع	٤١,٣	٢٢,٩	٣٨,١	٢١,٠	٤٩,٠	١٧,٤

* المصدر : إحصاء التوظيف ومستويات الأجور فى المؤسسات الخاصة، مصلحة الإحصاءات العامة ـ وزارة المالية والاقتصاد الوطنى ـ مطابع الشرق الأوسط ـ رجب ١٤٠٠ هـ، صفحة (١٩).

أمـا البيانـات الكمية فهى التى يمكن تقسيمها إلى مجموعات أو فئات متقاربة القيم. والبيانات الكمية تكون أحد نوعين، هما : البيانات المتصلة ، أو المستمرة (Continuous) والبيانات المنفصلة أو الوثابة (Discrete). فالبيانات الوثابة هى التى قد تختلف الوثبة فيها بين كل قراءة وأخرى، وهى لا تأخذ إلا قيماً معينة فقط، ولا يسمح لها بأخذ أى قيمة جديدة، بخـلاف القيم المعينـة. وتكون البيانات الوثابة عادة عبارة عن أعداد صحيحة، مثل عدد الأفراد، أو المؤسسات، أو الكتب، أو السيارات، أو الألوان.

وأما البيانات المستمرة فهى التى يمكن أن تأخذ أى قيمة خلال أى فترة أو مدى، أى لا توجد قيود عليها، ومثال ذلك البيانات الخاصة بالزمن، أو الطول، أو الوزن. ولتوضيح الفرق بين البيانات الوثابة والمستمرة يمكن اعتبار ظاهرة إقلاع أو هبوط الطائرات بأنها أحداث وثابة خلال فترة زمنية متصلة.

كذلك تعتبر الأهداف أثناء المباراة، أو حوادث المرور، كلها ظواهر وثابة أثناء فترات زمنية متصلة. ويعتبر التمييز بين البيانات الوثابة والمستمرة أمراً هاماً عند معالجة أى بيانات.

هذا وتنقسم البيانات المستمرة والوثابة إلى قسمين آخرين، وهما : البيانات المرتبة (Or- dered) والبيانات غير المرتبة، فالبيانات الخاصة بالأسعار لسلعة معينة، أو كميات المخزون بين يوم وآخر، أو درجات الحرارة أو البيانات الخاصة بمراقبة الإنتاج الصناعى، تعتبر من البيانات المرتبة لأن الترتيب الزمنى هام، لأنه يحتوى على معلومات أساسية لا غنى عنها.

ومن جهة أخرى، فإن بيانات العينة الخاصة بأعمار أو أوزان أو درجات الامتحانات لبعض الطلاب لا تحتاج لترتيب؛ إذ ليس من الضرورى اختيار الأشخاص بترتيب معين. ومثل هذه البيانات تحتاج لمجهود أكبر لتنظيمها، وهذا ما سوف يتم عرضه فى الأجزاء اللاحقة من هذا الفصل.

٣ - تبويب البيانات الوصفية البسيطة :

عرفت البيانـات الـوصفية على أنها البيانـات التى يمكن تصنيفها تحت صفات معينة، كالجنسيات أو المهن. لذلـك فإن أول عمـل يتم لتبويب البيانات الوصفية هو حصر تلك الصفات فى شكل عمودى.

يبدأ تفريغ البيانات بتسجيل العلامات، والعلامة هى خط عمودى لكل حالة، وتحزم كل أربع علامات (حالات) بالعلامة التى تمثل الحالة الخامسة أفقياً، وذلك لتسهيل عملية العد اليدوى. أما العمود الثالث المبين فى المثال التالى فهو الخاص بتسجيل أعداد العلامات.فهى إذاً عدد تكرار الحالات لكل صفة. لذلك فهى تسمى التكرارات ($ك_ر$)، فتكرار المجموعة

هو عدد المتغيرات فى نفس المجموعة، أما النمط الذى تم بموجبه توزيع التكرارات فيطلق عليه اسم التوزيع التكرارى.

مثال (٢ , ٣) :

البيانات التالية توضح مدى اعتماد الرؤساء فى ٣٠ وحدة تخطيطية* على مرؤوسيهم. ولقد كانت الإجابات على النحو الآتى :

كثيراً	كثيراً	أحياناً
أحياناً	كثيراً	كثيراً
كثيراً	أحياناً	كثيراً
نادراً	كثيراً	نادراً
كثيراً	أحياناً	أحياناً
نادراً	كثيراً	كثيراً
كثيراً	أحياناً	كثيراً
أحياناً	كثيراً	كثيراً
كثيراً	كثيراً	كثيراً
كثيراً	كثيراً	أحياناً

يبدأ التبويب بحصر الصفات فى العمود الأول، كما هو موضح فيما يلى ، يليه الحزم ثم التكرار.

درجة الاعتماد (الصفة)	التفريغ (الحزم)	التكرار (ك ر)
كثيراً	卌 卌 卌 ||||	١٩
أحياناً	卌 |||	٨
نادراً	|||	٣
المجموع		٣٠

*المصدر : د. على عبدالحفيظ : دور وحدات التخطيط فى الأجهزة الحكومية فى المملكة العربية السعودية ـ معهد الإدارة العامة ـ الرياض ـ ١٤٠٤ هـ صفحة (٧٤).

يعتبر المجموع للتكرارات هو أول الخطوات الهامة التى تلى عملية الجدولة، وذلك بهدف التأكد من أن هذا المجموع يساوى عدد المتغيرات، ولاستخداماته لاستخراج بعض المؤشرات الأساسية التى سيرد ذكرها فيما بعد، هذا ويمكن اعتبار جدول التوزيع التكرارى هو المكون من العمود الأول والعمود الأخير الخاص بالتكرارات.

معنى ذلك أن الجدول التكرارى البسيط يتكون من عدد من الصفوف (ص) مساوٍ لعدد الصفات وعمود واحد يمثل التكرارات، لذا يمكن كتابته على النحو (ص × ١).

أما إذا كان عدد الأعمدة (ع) أكثر من واحد، فيكون الجدول على النحو ص × ع وهو ما يسمى بالجدول التكرارى المزدوج، وأصغر جدول مزدوج هو ٢ × ٢ وهو ما يسمى بجدول الاقتران. والمثال التالى يوضح جدولاً للاقتران.

مثال(٣,٣) : جدول الاقتران

الحالة التعليمية / الجنس	متعلم	أمى	المجموع
سيدات	١٩	١١	٣٠
رجال	٢٧	٤	٣١
المجموع	٤٦	١٥	٦١

ويسمى المجموع الأفقى مع الحالة التعليمية بالتوزيع الهامشى (Marginal Distribution) للحالة التعليمية، بينما يسمى المجموع الرأسى مع الجنس بالتوزيع الهامشى للجنس.

٤ ـ تبويب البيانات الكمية :

تم اختيار عينة عشوائية من مرتكبى حوادث المرور فى إحدى الفترات فكانت أعمار أفراد العينة البالغ عددهم ١٥٠ فرداً على النحو التالى :

٣١	٣٠	٤١	٣٨	٣٢	٣١	٣١	٢٦
٢٩	٣١	٣١	٢٢	٣١	٣٨	٣٢	٣٣
٣٦	٣١	٣١	٣٥	٣١	٣٣	٤٢	٣١
٤٠	٣٣	٣٣	٣٠	٣٤	٣١	٣٥	٣٣
٢٦	٢٠	٣٢	٣٠	٣٢	٣٣	٢٦	٣١
٣٤	٢٩	٣٢	١٥	٢٦	٢٣	٣١	٣٣
٢٦	٢٨	٢٨	٣٢	٣٩	٣٠	٣٥	٢٢
٣١	٢٨	٣٦	٣١	٤٢	٣٢	٥٠	٣٥
٣١	٣٢	٣٢	٣١	٣٥	٣٥	٢٩	٣١
٣١	٤١	٢٥	٣٨	٣٠	٣٥	٣٣	٣٨
٣٦	٣٥	٣١	٣١	١٣	٣٥	٣٢	
٢٦	٣٣	٣٨	٣٣	٣٣	٢٩	٤٦	
٣١	٢٩	٤٤	٣١	١٩	٢٦	٣٤	
٢٢	٣٠	٢٧	٣٤	٣٤	٢٨	٣٠	
١٦	٤١	٢٦	٣٢	٣٦	٣٤	٢٣	
٣٣	١٨	٣١	٣٠	٣٢	٣٦	٣٤	
٣١	٣٨	٢٦	١٩	٣١	٣٣	٢٩	
٣٥	٣٥	٢٨	١٩	٢٥	٣٢	٢٧	
٣١	٣٤	٣٢	٤٤	٣١	٣٢	٣٠	
٣٣	٣٠	٢٧	٣٢	٥٠	٣١	١٦	

ولعله من الصعب جداً تحديد أهم خصائص توزيع هذه البيانات بدقة بمجرد النظر إليها. وبما أن الإحصاء علم يهتم بالمجموعات أكثر من اهتمامه بالمفردات، لذا فمن الجائز تلخيص هذه البيانات فى مجموعات متقاربة؛ ليتم عرضها بيانياً، ولتوضيح أهم خواص توزيعها الإحصائى، كذلك يمكن استخدام البيانات الملخصة (المبوبة) لتقدير أهم المؤشرات الإحصائية. إذاً فالهدف العام من تلخيص البيانات فى مجموعات هو استخراج المعلومات الأساسية الخاصة بالتوزيع الإحصائى للبيانات.

إن أول خطوة لتحديد المجموعات هى إيجاد المدى الذى يعرف بأنه الفرق بين أكبر قيمة وأصغر قيمة. يقسم المدى إلى أقسام تكون فى أكثر الحالات متساوية تسمى المجموعات أو

الفئات، ويسمى ناتج قسمة المدى (م) على عدد الفئات (ع) بطول الفئة (ط). وطول الفئة هو طول الخطوة التى يجب أن تخطوها من بداية الفئة (المجموعة) حتى نهايتها.

أى أن :

$$ط = \frac{م + ١}{ع}$$

وقد استبدلت قيمة م بالقيمة (م + ١)؛ لأن المدى الحقيقى أكبر من (م) بزيادة واحدة، لأنه يضم أكبر وأصغر قيمة معاً.

دعنا نتوقف الآن قليلاً لنتأمل البرنامج التالى، والذى يقوم بتحديد القيمة العظمى، والقيمة الصغرى لمجموعة من القيم.

```
10  REM برنامج لتحديد القيمه العظمى والقيمه الصغرى من مجموعه قيم
20  READ N   REM عدد القيم
30  READ X
40  PRINT , 'القيم'
50  PRINT ; '______'
60  PRINT
70  PRINT , X
80  H=X  REM القيمه العظمى
90  L=X  REM القيمه الصغرى
100 FOR I=1 TO N-1
110  READ X
120  PRINT , X
130  IF X>H THEN H=X
140  IF X<L THEN L=X
150 NEXT I
160 PRINT
170 PRINT ,H;'القيمه العظمى='
180 PRINT ,L;'القيمه الصغرى='
190 DATA 20,7,48,66,3,9,5,15,82,4,70,27,49,85,45,7,26,27,53,66,30
200 END
```

المخرجات

```
القيم
______

7
48
66
3
9
5
15
82
4
70
27
49
85
45
7
26
27
53
66
30

85    القيمه العظمى=
3     القيمه الصغرى=
```

لاستخدام هذا البرنامج لمجموعة مختلفة من البيانات، فإنك فقط تتأكد أن عدد القيم هو نفسه الموجود فى المتغير N فى البرنامج، وإلا فإنك تدخل القيمة التى تريدها. وبعد ذلك بالطبع لابد من تغيير البيانات الموجودة فى عبارة DATA بالبيانات الجديدة.

(٤ - ١): تحديد عدد الفئات وأطوالها

عند تكوين الجداول التكرارية فى جميع الحالات لكى تحقق البيانات المبوبة الهدف الذى أنشئت من أجله، يجب ملاحظة ما يلى:

١ - ألا يكون عدد الفئات كبيراً، فتقترب البيانات المبوبة من المفردات ويفقد التلخيص أهميته. كذلك يجب ألا يكون عدد الفئات صغيراً، فيفقد التوزيع التكرارى الكثير من تفاصيله الهامة، وذلك بوضع قيم متباعدة (متباينة) فى مجموعة واحدة.
وبما أن طول الفئة يزداد بنقصان عدد الفئات، فقد أثبتت الحالات التطبيقية أن أفضل عدد (ع) للفئات هو الذى يتراوح بين ١٠ و ١٥ فئة (١٠ ≦ ع ≦ ١٥). وقد يزيد أو ينقص بخمس فئات إذا دعت الضرورة. ويعتبر عدد الفئات صغيراً جداً إذا قل عن خمس فئات، وكبيراً جداً إذا زاد على العشرين فئة.

٢ - العدد الفردى للفئات أفضل من العدد الزوجى، أما الطول الزوجى للفئات فأفضل من الفردى.

٣ - يجب تحديد الفئات بحيث تكون البيانات الخام متمركزة حول نقطة الوسط لكل فئة. ونقطة الوسط تسمى مركز الفئة، وهى عبارة عن متوسط الحد الأدنى والحد الأعلى لكل فئة.

٤ - يفضل أن يكون الحد الأدنى، وكذلك الحد الأعلى، للفئة عدداً صحيحاً خالياً من الكسور، إذا كان الطول زوجياً. أما إذا كان الطول فردياً، فيفضل أن يحتوى كل حد على ٠,٥ وذلك بهدف جعل مراكز الفئات أعداداً صحيحة خالية من الكسور فى جميع الحالات.

٥ - يفضل أن تكون أطوال الفئات متساوية، فالأطوال غير المتساوية تعوق عمليتى العرض البيانى للتوزيعات، واستخراج بعض المؤشرات ومقارنة متغيرات المجموعات المختلفة. إلا أن ذلك قد تحتمه بعض الظروف العملية، خاصة إذا احتوت البيانات على بعض الفراغات.

٦ - يفضل عدم استخدام الفئات المفتوحة، وهى الفئات التى لا يكون لها حد أدنى للفئة الأولى، أو حد أعلى للفئة الأخيرة؛ ذلك لأن هذه الفئات تجعل مهمة العرض البيانى

أكثر صعوبة، وتقلل من عدد المؤشرات الهامة. بل إن أهم المؤشرات لا يمكن استخراجها، ما دامت هناك بعض الفئات المفتوحة، إلا أن الضرورة قد تفرض استخدام هذا النوع من الفئات، فالبيانات الخاصة بالأعمار أو الدخل يصعب تحديد سقف لها فى بعض الحالات.

(٤ - ٢) طريقة ستيرقس (Sturges)

تتلخص طريقة ستيرقس لتحديد أنسب طول للفئة (ط) فى المعادلة التالية :

$$ط = \frac{\text{المدى}}{١ + ٣,٣٢٢ \times \text{لو}_{١٠} \text{ن}}$$

حيث ن هى عدد المتغيرات المعنية بالدراسة، أى مجموع التكرارات. فلتحديد طول الفئة للبيانات السالفة الذكر والخاصة بالأعمار يتضح أن :

$$ط = \frac{٥٠ - ١٣}{١ + ٣,٣٢٢ \times \text{لو} ١٥٠}$$

$$ط = ٤,٥$$

أى أن طول الفئة ٤ أو ٥. فإذا كان طول الفئة ٤ فعدد الفئات

$$ع = \frac{٣٧}{٤}$$

$$\therefore ع = ١٠$$

أما إذا كان الطول ٥ فعدد الفئات يساوى

$$\frac{٣٧}{٥} = ٨ \text{ تقريباً.}$$

إلا أن نظرية ستيرقس ليست ملزمة فى جميع الحالات، ولكنها تعطى مؤشراً مناسباً لطول الفئة، وعدد الفئات فى كل حالة. ويلاحظ هنا أن حاصل ضرب عدد الفئات فى الطول يكون أكبر من المدى؛ لذا يمكن تسميته بالمدى النظرى. فالمدى النظرى فى الحالة الأولى يساوى ٤٠، وكذلك فى الحالة الثانية (٥ × ٨)، لذلك يفضل اختيار أصغر مدى نظرى ممكن؛ لتقليل الفرق (الفراغ) بقدر الإمكان.

فيما يلي برنامج لحساب أنسب طول للفئة حسب طريقة ستيرقس التى سبق شرحها، وهنا نستخدم البرنامج السابق لتحديد القيمة العظمى والقيمة الصغرى، ثم نحسب المدى ومن ثم طول الفئة.

لاستخدام هذا البرنامج لبيانات مختلفة، فإنك كذلك تغير قيمة N بالإضافة إلى البيانات نفسها فى عبارة DATA .

وأما المعادلة المستخدمة هنا فهى :

$$C = \frac{R}{[1 + 3.322 \ \text{LOG}(N)]}$$

حيث :

الطول المناسب للفئة = C

المدى = R

عدد المتغيرات = N

```
10  REM   برنامج لحساب انسب طول للفئه وعدد الفئات حسب طريقه ستيرقس
20  READ N    REM      عدد القيم
30  READ X
60  PRINT
80  H=X  REM القيمه العظمى
90  L=X  REM القيمه الصغرى
100 FOR I=1 TO N-1
110  READ X
130  IF X>H THEN H=X
140  IF X<L THEN L=X
150 NEXT I
160 R=H-L  REM  المدى
170 C=R/(1+3.322*LGT(N))
180 PRINT
190 PRINT ,H; '= القيمه العظمى'
195 PRINT
200 PRINT ,L; '= القيمه الصغرى'
205 PRINT
210 PRINT ,R; '= المدى'
220 PRINT
230 F=INT(C+0.5)
240 G=INT(R/F+1)
250 PRINT ,F; '= طول الفئه'
260 PRINT
270 PRINT ,G; '= عدد الفئات'
280 DATA 150
290 DATA 31,29,36,40,26,34,26,31,31,31,36,26,31,22,16,33,31,35,31,33
300 DATA 30,31,31,33,20,29,28,28,32,41,35,33,29,30,41,18,38,35,34,30
310 DATA 41,31,31,33,32,32,28,36,32,25,31,38,44,27,26,31,26,28,32,27
320 DATA 38,22,35,30,30,15,32,31,31,38,31,33,31,34,32,30,19,19,44,32
330 DATA 32,31,31,34,32,26,39,42,35,30,13,33,19,34,36,32,31,25,31,50
340 DATA 31,38,33,31,33,23,30,32,35,35,35,29,26,28,34,36,33,32,32,31
350 DATA 31,32,42,35,26,31,35,50,29,33,32,46,34,30,23,34,29,27,30,16
360 DATA 26,33,31,33,31,33,22,35,31,38
370 END
```

المخرجات

```
القيمه العظمى =   50
القيمه الصغرى =   13
المدى =   37
طول الفئه =   4
عدد الفئات =   10
```

مثال (٤ , ٣) :
كون جدول توزيع تكرارى للبيانات الخاصة بأعمار مرتكبى حوادث المرور خلال إحدى الفترات الزمنية .

الحل :

(١) المدى = ٥٠ - ١٣
= ٣٧

(٢) ورد من قبل أن أفضل عدد للفئات يتراوح بين ١٠ ـــ ١٥ فئة .

والجدول التالى يبين الطول، والمدى النظرى فى كل حالة .

عدد الفئات (ع)	طول الفئة ط = ٣٧ ÷ ع	المدى النظرى ع × ط
١٠	٤	٤٠
١١	٤	٤٤
١٢	٤	٤٨
١٣	٣	٣٩
١٤	٣	٤٢
١٥	٣	٤٥

يتضح من الجدول السابق أن أصغر مدى نظرى هو ٣٩ وهو يقابل ١٣ فئة، طول كل منها ٣. ولا يعنى ذلك أن هذا هو أفضل اختيار، فلقد اتضح من طريقة ستيرقس أنه بالإمكان استخدام ١٠ فئات طول كل منها ٤، أو ٨ فئات طول كل منها ٥. كما أن النواحى العملية قد تغلب على النظرية فيتحدد طول الفئة وعدد الفئات مسبقاً.
يتميز اختيار ١٣ فئة بأنه عدد فردى، إلا أن الطول الفردى غير مرغوب فيه كما ورد من قبل. أما الفراغ وهو الفرق بين المدى النظرى والمدى (٣٩ - ٣٧ = ٢) فيجب أن يترك كله أو جزء منه للفئة الأخيرة. لذا فقد وقع الاختيار هنا على ترك ٠,٥ للبداية والباقى (١,٥) للفئة الأخيرة. ولقد تم تقسيم الفراغ على هذا النحو بسبب الطول الفردى (٣) لتبدأ الفئة الأولى من ١٣ - ٠,٥ = ١٢,٥

وتنتهى الفئة الأخيرة فى :
٥٠ + ١,٥ = ٥١,٥ .

والجدول التالى يوضح الفئات والعلامات والتكرارات :

الفئة	العلامات (التفريغ)	التكرار (ك ر)
١٢,٥ - ١٥,٥	ǀǀ	٢
١٥,٥ - ١٨,٥	ǀǀǀ	٣
١٨,٥ - ٢١و٥	ǀǀǀǀ	٤
٢١,٥ - ٢٤,٥	卌	٥
٢٤,٥ - ٢٧,٥	卌 卌 ǀǀǀǀ	١٤
٢٧,٥ - ٣٠,٥	卌 卌 卌 卌 ǀ	٢١
٣٠,٥ - ٣٣,٥	卌 卌 卌 卌 卌 卌 卌 卌 卌 卌 卌 ǀǀǀǀ	٥٩
٣٣,٥ - ٣٦,٥	卌 卌 卌 卌 ǀǀǀǀ	٢٤
٣٦,٥ - ٣٩,٥	卌 ǀǀ	٧
٣٩,٥ - ٤٢,٥	卌	٦
٤٢,٥ - ٤٥,٥	ǀǀ	٢
٤٥,٥ - ٤٨,٥	ǀ	١
٤٨,٥ - ٥١,٥	ǀǀ	٢
المجموع		١٥٠

البرنامج التالى يقوم بتفريغ وتبويب البيانات بافتراض أن الفئات معلومة، علماً بأنه سبق إعداد برنامج لتحديد القيمة الصغرى والقيمة العظمى والمدى، حيث يمكن استخدام ذلك البرنامج أو البرنامج الخاص بطريقة ستيرقس، لتحديد الفئات :

```
10  REM        برنامج لتفريغ وتبويب البيانات
20  DIM X(150),F(13),L(13),U(13)
30  READ N1,N2      REM N1=NO OF OBSERVATIONS, N2=NO OF CLASSES
35  DATA 150,13
40  FOR I=1 TO N1
50   READ X(I)
60  NEXT I
70  FOR I= 1 TO N2
80   READ L(I),U(I)
90  NEXT I
100 FOR I= 1 TO N1
105  FOR J= 1 TO N2
110   IF X(I)>=L(J) .AND. X(I)<U(J) THEN F(J)=F(J)+1
235  NEXT J
240 NEXT I
245 PRINT USING 250
246 PRINT USING 260
247 PRINT USING 270
250 :         ________________________________
260 :          التكرار              الفئه
270 :         ________________________________
280 :            ##              ##.# – ##.#
290 FOR I=1 TO N2
300  PRINT USING 280,F(I),U(I),L(I)
310 NEXT I
320 DATA 31,29,36,40,26,34,26,31,31,31,36,26,31,22,16,33,31,35,31,33
330 DATA 30,31,31,33,20,29,28,28,32,41,35,33,29,30,41,18,38,35,34,30
340 DATA 41,31,31,33,32,32,28,36,32,25,31,38,44,27,26,31,26,28,32,27
350 DATA 38,22,35,30,30,15,32,31,31,38,31,33,31,34,32,30,19,19,44,32
360 DATA 32,31,31,34,32,26,39,42,35,30,13,33,19,34,36,32,31,25,31,50
370 DATA 31,38,33,31,33,23,30,32,35,35,35,29,26,28,34,36,33,32,32,31
380 DATA 31,32,42,35,26,31,35,50,29,33,32,46,34,30,23,34,29,27,30,16
390 DATA 26,33,31,33,31,33,22,35,31,38
400 DATA 12.5,15.5,15.5,18.5,18.5,21.5,21.5,24.5
410 DATA 24.5,27.5,27.5,30.5,30.5,33.5,33.5,36.5
420 DATA 36.5,39.5,39.5,42.5,42.5,45.5,45.5,48.5,48.5,51.5
430 END
```

المخرجات

التكرار	الفئه
2	15.5 – 12.5
3	18.5 – 15.5
4	21.5 – 18.5
5	24.5 – 21.5
14	27.5 – 24.5
21	30.5 – 27.5
59	33.5 – 30.5
24	36.5 – 33.5
7	39.5 – 36.5
6	42.5 – 39.5
2	45.5 – 42.5
1	48.5 – 45.5
2	51.5 – 48.5

التوزيع التكرارى السالف الذكر هو توزيع تكرارى مستمر، لذا فإن كل فئة تبدأ من حدها الأدنى إلى ما قبل حدها الأعلى، أى أن الحد الأعلى لكل فئة يتبع للفئة التالية، حيث يكون حداً أدنى. أما فى حالة التوزيع التكرارى الوثاب فيكون الحد الأعلى للفئة مختلفاً عن الحد الأدنى للفئة التالية، كما هو مبين فيما يلى :

مثال (٣،٥) :

توزيع تكرارى وثاب

الفئة	التكرار (ك ر)
٣ — ٧	١
٨ — ١٢	٣
١٣ — ١٧	٦
١٨ — ٢٢	٤
٢٥ — ٢٩	١
المجموع	١٥

يعرف طول الفئة فى حالة المجموعات المستمرة بأنه الفرق بين الحد الأعلى والحد الأدنى لنفس الفئة ، بينما يعرف بأنه نفس ذلك الفرق مضافاً إليه واحد فى حالة الجداول الوثابة ، أما مركز الفئة فهو الحد الأدنى للفئة مضافاً إليه نصف طولها . وتمثل مراكز الفئات متوالية حسابية فرقها (أساسها) مساوٍ لطول الفئة فى حالة الأطوال المتساوية للفئات المستمرة، أو الوثابة ذات الوثبات الثابتة .

مثال (٣،٦) :

استخدم الجدول التكرارى التالى لإيجاد مراكز الفئات ، وناتج ضرب كل مركز فئة فى التكرار المقابل، ومجموع التكرارات ، ومجموع نواتج الضرب . والجدول هو :

المجموعات	التكرار (ك ر)
٦ — ١٠	٣
١٠ — ١٤	٧
١٤ — ١٨	١٥
١٨ — ٢٢	١٠
٢٢ — ٢٦	١

الحل :

المجموعات	$ك_ر$	مراكز الفئات $س_ر$	ناتج الضرب $س_ر × ك_ر$
٦ ـ ١٠	٣	٨	٠٢٤
١٠ ـ ١٤	٧	١٢	٠٨٤
١٤ ـ ١٨	١٥	١٦	٢٤٠
١٨ ـ ٢٢	١٠	٢٠	٢٠٠
٢٢ ـ ٢٦	١	٢٤	٠٢٤
المجموع ($\sum$)	٣٦		٥٧٢

٥ ـ التجمع التكرارى :

يحتاج الشخص أحياناً لتجميع التكرارات تجميعاً تراكمياً من طرف إلى آخر؛ للإجابة على أحد السؤالين التاليين :

- كم عدد المتغيرات التى تقل عن الحد الأعلى لفئة معينة؟
- كم عدد المتغيرات التى تزيد على الحد الأدنى لفئة معينة؟

ويستخدم التجمع التكرارى الصاعد للاجابة عن السؤال الأول ، بينما يستخدم التجمع التكرارى النازل (الهابط) للاجابة عن السؤال الثانى . ويمكن الحصول على التجمع الصاعد إذا جمعت التكرارات من الفئة الأولى (الصغرى) إلى الفئة الأخيرة (العليا)، بينما تبدأ عملية الجمع التراكمى بطريقة عكسية فى حالة التجمع التكرارى الهابط .

مثال (٣,٧) :

استخدم الجدول التكرارى الخاص بالمثال (٤) لتكوين جدول تجمع صاعد وجدول تجمع هابط ، ثم أوجد الفرق بين التجمع الصاعد والهابط عند كل فئة ، وأوجد عدد مرتكبى الحوادث الذين تزيد أعمارهم على ٣٦,٥ سنة ، وكذلك عدد الذين تقل أعمارهم عن ٣٦,٥ سنة .

الحل :

الفئة	التكرار (ك ر)	التجمع التكرارى الصاعد (أقل من الحد الأعلى)	التجمع التكرارى الهابط (أكثر من الحد الأدنى)	الفرق الصاعد ـ الهابط
١٥,٥ـ١٢,٥	٢	٢	١٥٠	١٤٨-
١٨,٥ـ١٥,٥	٣	٥	١٤٨	١٤٣-
٢١,٥ـ١٨,٥	٤	٩	١٤٥	١٣٦-
٢٤,٥ـ٢١,٥	٥	١٤	١٤١	١٢٧-
٢٧,٥ـ٢٤,٥	١٤	٢٨	١٣٦	١٠٨-
٣٠,٥ـ٢٧,٥	٢١	٤٩	١٢٢	٧٣-
٣٣,٥ـ٣٠,٥	٥٩	١٠٨	١٠١	٧
٣٦,٥ـ٣٣,٥	٢٤	١٣٢	٤٢	٩٠
٣٩,٥ـ٣٦,٥	٧	١٣٩	١٨	١٢١
٤٢,٥ـ٣٩,٥	٦	١٤٥	١١	١٣٤
٤٥,٥ـ٤٢,٥	٢	١٤٧	٥	١٤٢
٤٨,٥ـ٤٥,٥	١	١٤٨	٣	١٤٥
٥١,٥ـ٤٨,٥	٢	١٥٠	٢	١٤٨
المجموع	١٥٠			

١ ـ عدد الذين تزيد أعمارهم على ٣٦,٥ = ١٨ شخصاً.
٢ ـ عدد الذين تقل أعمارهم عن ٣٦,٥ = ١٣٢ شخصاً.

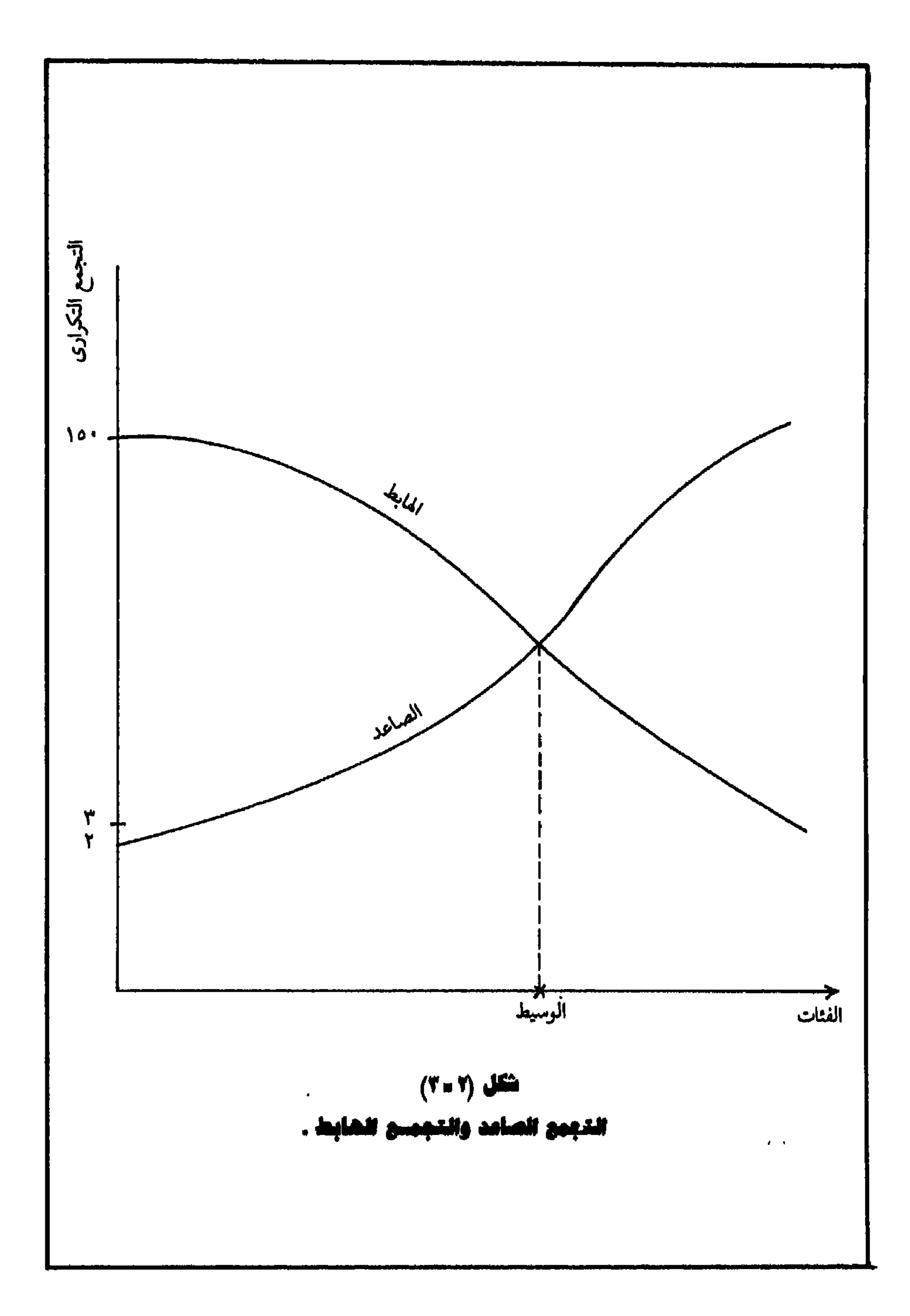
تسمى نقطة تقاطع التجمع التكرارى الصاعد مع التجمع التكرارى الهابط بالوسيط. هذا ويكون الوسيط فى الفئة التى يكون الفرق بين التجمعين عندها أقل ما يمكن، لذلك تسمى هذه الفئة بالفئة الوسيطية. إذاً فالفئة ٣٠,٥ ـ ٣٣,٥ هى الفئة الوسيطية.

البرنامج التالى يقوم بحساب التجمع التكرارى الصاعد والتجمع التكرارى الهابط، وقد صمم البرنامج، شأنه شأن كل البرامج، بطريقة عامة، أى أن البرنامج يصلح لأى مجموعة مختلفة من البيانات فقط تتغير قيمة N والبيانات فى عبارة DATA .

```
10   REM     برنامج لحساب التجمع التكراري الصاعد والهابط
20   DIM A(13),B(13),C(13),D(13),E(13),F(13)
30   PRINT USING 250
40   PRINT USING 230
50   PRINT USING 240
60   PRINT USING 250
70   PRINT
75   F1=0
80   READ N      REM  NO OF OBSERVATIONS
90   FOR I=1 TO N
100   READ A(I),B(I),F(I)   REM الحد الادنى، الحد الاعلى والتكرار
110   F1=F1+F(I)
120  NEXT I
130  C(1)=F(1)
140  D(1)=F1
150  FOR I=2 TO N
160   C(I)=C(I-1)+F(I)
170   D(I)=D(I-1)-F(I-1)
180  NEXT I
190  FOR I=1 TO N
200   E(I)=C(I)-D(I)
210   PRINT USING 260, E(I),D(I),C(I),F(I),B(I),A(I)
220  NEXT I
222  PRINT
225  PRINT
230  :     الفرق    تجمع تكراري    تجمع تكراري    التكرار       الفئة
240  :              هابط            صاعد
250  :  ____________________________________________________________
260  :      ####        ####        ####        ###      ##.# - ##.#
270  DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
280  DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
290  DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
300  END
```

المخرجات

الفرق	تجمع تكراري هابط	تجمع تكراري صاعد	التكرار	الفئة
-148	150	2	2	15.5 - 12.5
-143	148	5	3	18.5 - 15.5
-136	145	9	4	21.5 - 18.5
-127	141	14	5	24.5 - 21.5
-108	136	28	14	27.5 - 24.5
-73	122	49	21	30.5 - 27.5
7	101	108	59	33.5 - 30.5
90	42	132	24	36.5 - 33.5
121	18	139	7	39.5 - 36.5
134	11	145	6	42.5 - 39.5
142	5	147	2	45.5 - 42.5
145	3	148	1	48.5 - 45.5
148	2	150	2	51.5 - 48.5



شكل (٢ - ٣)

التجمع الصاعد والتجمع الهابط .

٦ ـ العرض البيانى :

يهدف العـرض البيانى إلى إعطاء صورة أوضح للتوزيع التكرارى للمتغيرات، ومن ثم معرفة الاتجاه العام للمتغيرات. وقد يكون الرسم البيانى خطياً أو عمودياً أو دائرياً أو مجسداً، وسوف يتم استعراض الأنواع الثلاثة الأولى لأنها الأكثر استخداماً.

(٦ـ١) الرسم البيانى الخطى :

يستخدم الرسم البيانى الخطى سواء على أوراق حسابية (Arithmetic) أو أوراق لوغاريتمية (Logarithmic) أو نسبية ، وذلك بهدف توضيح التغيرات المطلقة (الزيادة أو النقصان)، ومن ثم تحديد الاتجاهات العامة. لذلك يكثر استخدام الرسم البيانى الخطى لقياس التغيرات من وقت إلى آخر (السلاسل الزمنية)، سواء فى مجالات الأسعار أو الدخل أو المبيعات، وهو أبسط وأسهل أنواع الرسم البيانى، ويمكن استخدامه للبيانات المستمرة والوثابة.

مثال (٣,٨) :

يوضح الرسم البيانى الخطى التالى التضخم (معدلات التغير السنوى فى معامل الانكماش) فى المملكة العربية السعودية خلال الفترة ١٣٩٤/١٣٩٥ ـ ١٤٠٣/ ١٤٠٤هـ.

شكل (٣ - ٣)

التضخم فى المملكة العربية السعودية

(معدلات التغير السنوى فى معامل انكماش الناتج المحلى الإجمالى غير النفطى)*

معدل التغير السنوى

٦٥
٦٠
٥٥
٥٠
٤٥
٤٠
٣٥
٣٠
٢٥
٢٠
١٥
١٠
٥
٠
٥ -
١٠ -

٩٠/١٣٨٩
٩١/١٣٩٠
٩٢/١٣٩١
٩٣/١٣٩٢
٩٤/١٣٩٣
٩٥/١٣٩٤
٩٦/١٣٩٥
٩٧/١٣٩٦
٩٨/١٣٩٧
٩٩/١٣٩٨
٤٠٠/١٣٩٩
٤٠١/١٤٠٠
٤٠٢/١٤٠١
٤٠٣/١٤٠٢
٤٠٤/١٤٠٣

السنوات

* المصدر : مؤسسة النقد العربى السعودى - التقرير السنوى لعام ١٤٠٤ هـ - الرياض ١٤٠٥ ، صفحة (٧).

(٦ ـ ٢) الرسم البياني العمودى :

وهـو إما أن يكون رأسياً، أو أفقياً، أو مزدوجاً. ويعتبر النوع الأول أكثر استخداماً من الأخـيرين، إذ تكون الفئات على المحور الأفقى (السينى)، وتكون التكرارات على المحور الرأسى (الصادى)، بينما تسجل مراكز الفئات بمراكز الأعمدة. هذا ويجب أن تكون الأعمدة متساوية العـرض مع ضرورة تساوى المسافات بينها. أما إذا اعتبرت التكرارات عبارة عن إحداثيات خاصة بمراكز الفئات، وتم توصيلها فيسمى الرسم البيانى الناتج بالمضلع التكرارى.

مثال (٩ , ٣) :

يوضـح الرسم البيانى التالى نصيب القطاع الخاص فى الإنتاج المحلى الإجمالى الحقيقى خلال الفترة ١٣٩٤/١٣٩٥ ـ ١٤٠٣/١٤٠٤ هـ*.

(*) المصدر : نفس المصدر السابق، صفحة ٢٣ .

شكل (٤ - ٣)

نصيب القطاع الخاص فى الانتاج المحلى الاجمالى الحقيقى

(٦ - ٣) الرسم البيانى الدائرى :

يستخدم الرسم البيانى الدائرى لتوضيح المقادير النسبية ، ويمثل العدد الكلى للدرجات (٣٦٠ درجة) بنسبة ١٠٠٪ ، ويحسب القطاع لكل متغير بنسبة مقداره إلى المقدار الكلى ، أى أن زاوية القطاع تساوى النسبة المئوية مضروبة فى ٣٦٠.

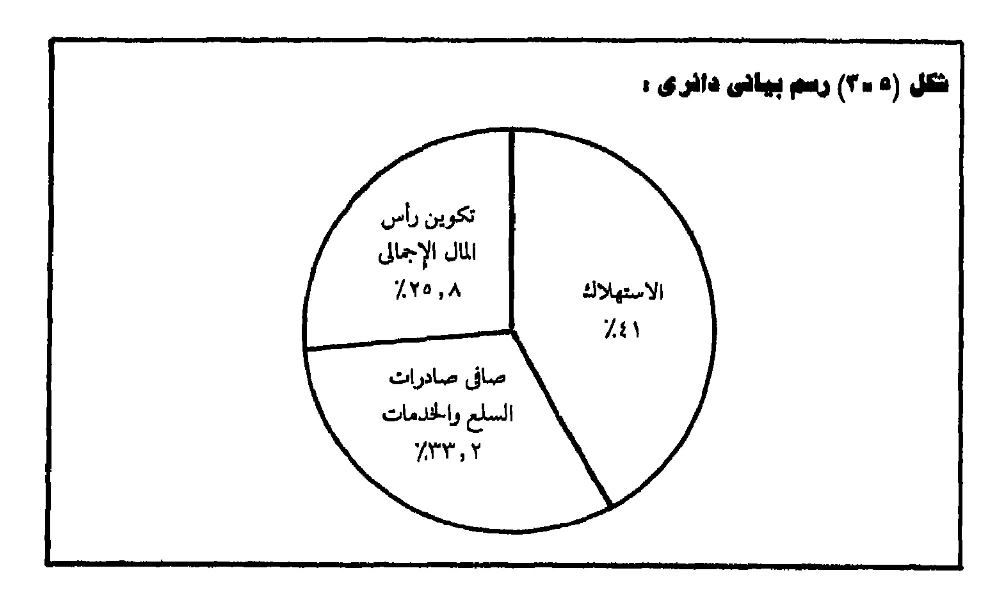
مثال (٣,١٠) :

يوضح الجدول التالى والرسم البيانى الدائرى الانفاق على الانتاج المحلى الاجمالى بالمملكة العربية السعودية خلال عام ١٣٩٩/١٤٠٠ هـ* .

المصدر	القيمة	النصيب المئوى
الاستهلاك	١٥٨٣٩٢	٤١
تكوين رأس المال الاجمالى	٩٩٨٤٦	٢٥,٨
صافى صادرات السلع والخدمات	١٢٨٢١٥	٣٣,٢
مجموع الانفاق المحلى	٣٨٦٤٥٣	١٠٠

* المصدر : مؤسسة النقد السعودى - التقرير السنوى لعام ١٤٠١ هـ (١٩٨١م) ، الرياض - صفر ١٤٠٢ هـ ، صفحة (١٩).

شكل (٣-٥) رسم بيانى دائرى :



تمارين

١ ـ عرِّف الفرق بين القيم العينية والمجتمع والعلاقة بينهما.
٢ ـ ما هو الفرق بين البيانات الوصفية والبيانات الكمية؟ وهل يمكن قياس جميع البيانات بنفس نوع القياس؟ ولماذا؟
٣ ـ حدد أنواع القياس التى يمكن استخدامها.
٤ ـ البيانات التالية توضح تقديرات ٥٠ دارساً، والمطلوب عرضها فى جدول تكرارى، والبيانات هى :
جيد، جيد جداً، جيد جداً، جيد، مقبول، جيد، جيد جداً، مقبول، ممتاز، جيد، جيد، جيد جداً، جيد، ممتاز، جيد، جيد جداً، راسب، جيد، جيد، جيد، جيد جداً، مقبول، راسب، جيد، مقبول، جيد، ممتاز، جيد، جيد جداً، راسب، جيد، مقبول، جيد، جيد جداً، راسب، مقبول، جيد، جيد جداً، راسب، جيد، مقبول، مقبول، جيد جداً، مقبول، جيد، جيد، جيد جداً، راسب، مقبول، جيد جداً.
ملحوظة : البيانات أعلاه تسمى بيانات مرحلية (Interval) .
٥ ـ استخدم البيانات أعلاه لإيجاد عدد الذين :
١ ـ نجحوا.
٢ ـ حصلوا على تقدير جيد على الأقل.
٣ ـ حصلوا على أقل من جيد جداً.
٤ ـ حصلوا على أقل من جيد.
٥ ـ حصلوا على تقدير أعلى من جيد.
٦ ـ رسبوا.
٧ ـ حصلوا على تقدير ممتاز .
٦ ـ سئل كل من مائة رياضى عن أحب أنواع الرياضة إلى نفسه، فقام الباحث بترميز الإجابات على النحو الآتى :
السباحة = ١
كرة القدم = ٢
كرة السلة = ٣

الكرة الطائرة = ٤
التنس = ٥
ألعاب القوى = ٦
فكانت النتائج على النحو الآتى :
٣، ٢، ٢، ٦، ٢، ٢، ٤، ١، ٢، ١، ٥، ٦، ٣، ٢، ٥، ١، ٣، ٤، ١، ٢، ٣، ٥،
٣، ٢، ١، ٤، ١، ٢، ٤، ٢، ٤، ٣، ٢، ١، ٦، ٥، ٦، ٢، ١، ٢، ٢، ٣، ٢، ١،
٤، ٢، ١، ٦، ٣، ٤، ٢، ٢، ٦، ٢، ٢، ٢، ٣، ١، ٦، ٤، ٢، ٣، ١، ٥، ٥، ٦،
٢، ٤، ٢، ٣، ١، ٢، ١، ١، ٢، ٣، ٤، ٢، ٥، ٢، ٦، ١، ١، ٢، ٤، ٣، ٣، ٤،
٢، ٢، ١، ٥، ٢، ٦، ٤، ٣، ١، ١، ٢، ١

البيانات أعلاه تسمى بيانات اسمية (Nominal) .
والمطلوب هو :
أ ـ إعداد جدول تكرارى.
ب ـ رسم بيانى دائرى.
جـ ـ رسم بيانى عمودى.

٧ ـ قامت إحدى شركات تسويق الشاى باستقصاء عينة من الزبائن قوامها ٦٠ شخصاً عن آرائهم فى مستوى أحد الأنواع على أن تكون الإجابة بين الصفر والأربعة، وذلك باعتبار أن الصفر يعنى أن مستوى الشاى سىء جداً، بينما تعنى الأربعة أن المستوى ممتاز، فكانت الإجابات على النحو الآتى :
٣، ٤، ٢، ٣، ١، ٢، ٣، ٠، ٤، ٣، ٢، ٣، ١، ١، ١، ٢، ٤، ٣، ٢، ٠، ٢، ٣، ٣، ٢،
٤، ١، ٠، ٣، ٠، ٣، ٢، ١، ٠، ٤، ٣، ٢، ٣، ٢، ٣، ٤، ٣، ٣، ٢، ١، ٢، ٤، ٣، ٢،
٣، ٠، ٠، ٣، ٢، ٤، ١، ١، ٣، ٢، ١، ٣
(ملحوظة : يسمى هذا النوع من البيانات بالبيانات التسلسلية [ORDINAL]) .
والمطلوب هو :
١ ـ إعداد جدول تكرارى.
٢ ـ عدد الذين كانت تقديراتهم ٣ فما فوق.
٣ ـ عدد الذين كانت تقديراتهم أقل من ٣ .
٤ ـ بكم تقدر عدد الذين يعتقدون أن مستوى هذا النوع من الشاى ممتاز، إذا كان عدد المستهلكين يساوى ٢ , ١ مليون شخص.

٥ ـ ارسم رسماً بيانياً عمودياً للبيانات .
٦ ـ ارسم رسماً بيانياً دائرياً للبيانات .

٨ ـ البيانات التالية توضح أسعار بعض المواد التى أخذت كعينة من أحد المحلات التجارية :

٥٣	١١٦	٧٠	٥٤	٣٦
٧٠	٩٨	١٧٦	١٣٩	٣٨
٦٧	١٠٢	١٦١	٧٧	١٣٧
٩٣	١٣٣	٥٦	٩٣	٨١
٨١	١٣١	١٨٥	٨٨	٥٤
٨٦	٩٢	٨٣	٦٠	١٠٩
٩٥	٢٤	١١٧	٣٥	٥٧
١٠٦	١٩	١٧١	٨١	٧٩
١٠٥	٥٩	٥٥	١١٠	١٣٦

(ملحوظة : البيانات أعلاه تسمى بيانات نسبية [RATIO])

المطلوب هو :

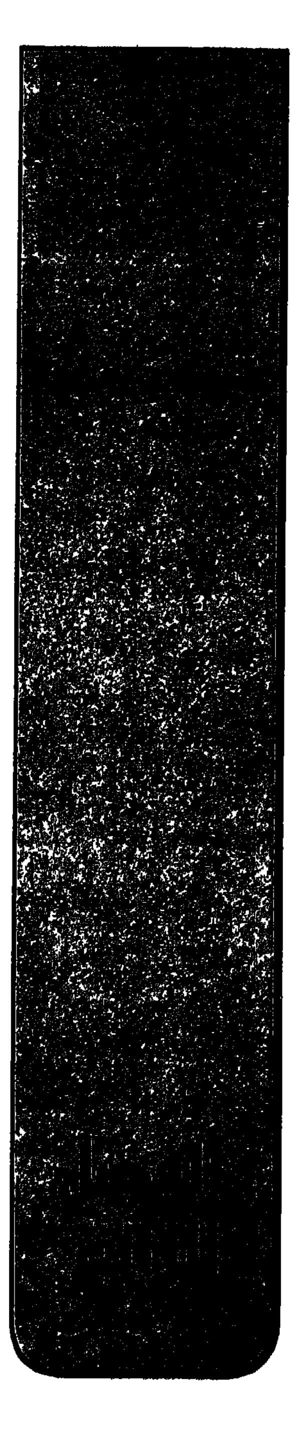
١ ـ المدى .
٢ ـ استخدم طريقة ستيرقس لتحديد عدد الفئات وطول كل فئة .
٣ ـ جدول تكرارى مناسب .
٤ ـ جدول تكرارى بحيث يكون الحد الأدنى للفئة الأولى صفراً والحد الأعلى للفئة الأخيرة ٢٠٠ ، وعدد الفئات ثمانياً .
٥ ـ استخدم الجدول التكرارى فى (٣) لإيجاد جدول متجمع صاعد وآخر هابط ، وحدد الفئة الوسيطية .
٦ ـ استخدم الجدول التكرارى فى (٤) لإيجاد جدول متجمع صاعد وآخر هابط وحدد الفئة الوسيطية .
٧ ـ قارن الجدول فى (٤) بالجدول فى (٣) ووضح أيهما أفضل .
٨ ـ ارسم الرسومات البيانية التالية :
أ ـ رسم بيانى خطى للبيانات .
ب ـ رسم بيانى عمودى .
جـ ـ رسم بيانى دائرى. وذلك اعتماداً على (٤) .

٩ ـ استخدم الأسئلة ٤ ، ٥ ، ٦ ، ٧ ، ٨ لتعريف ما يأتى :
أ ـ البيانات الاسمية .
ب ـ البيانات التسلسلية .
جـ ـ البيانات المرحلية .
د ـ البيانات النسبية .

١٠ ـ باستخدام البيانات الواردة فى السؤال (٨) اكتب برنامج بيسك لحساب الآتى :
١) القيمة العظمى .
٢) القيمة الصغرى .
٣) المدى .
٤) أنسب طول للفئة حسب طريقة ستيرقس .

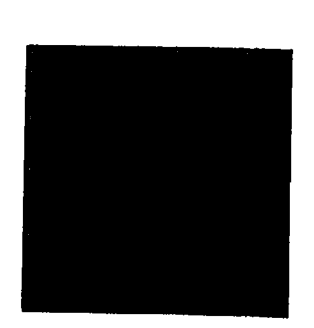
١١ ـ استخـدم بيانات الجدول التكرارى فى السؤال (٨) الجزء (٣) واكتب برنامج بيسك لإيجاد التجمع التكرارى الصاعد ، والتجمع التكرارى الهابط ، والفئة الوسيطية .

١٢ ـ اكتب برنامجاً متكاملاً لبيانات السؤال (٨) لإيجاد ما يلى :
أ ـ أفضل عدد للفئات وطول الفئة اعتماداً على أصغر مدى نظرى .
ب ـ تبويب وتفريغ البيانات اعتماداً على أفضل طول وعدد للفئات .



مقاييس النزعة المركزية





مقاييس النزعة المركزية

Measures of Central Tendency

١ - الإحصائية (Statistic) :

تعرف الإحصائية بأنها دالة العناصر (الوحدات) المكونة للعينة والتى لا تعتمد على معالم مجهولة. فالإحصائيات (Statistics) هى كميات يمكن حسابها من واقع قيم المفردات المكونة للعينة التى تم اختيارها، لذا فهى تكون الأساس الذى يرتكز عليه الاستنتاج الإحصائى أو الوصف الإحصائى للبيانات.

ومقاييس النزعة المركزية هى المجموعة الأولى من الإحصائيات الخاصة بتحديد معالم التوزيع التكرارى للبيانات المبوبة أو المفردات، وسميت هذه المجموعة بمقاييس النزعة المركزية، لأنها تهدف إلى تحديد نقطة معينة تتجمع حولها بقية القيم. وتعتبر هذه النقطة أفضل ممثل لمجموعة البيانات التى حُسبت منها؛ لأنها تحمل من صفات المجتمع أكثر مما تحمله أى نقطة أخرى.

هذا وتنقسم مقاييس النزعة المركزية إلى خمسة أنواع، هى :

١ - الوسط الحسابى	Arithmetic Mean
٢ - الوسيط	Median
٣ - المنوال	Mode
٤ - الوسط الهندسى	Geometric Mean
٥ - الوسط التوافقى	Harmonic Mean

٢ ـ الوسط الحسابى ($\bar{س}$) :

إذا كانت $س_١$ ، $س_٢$ ، $س_٣$، . . . . . . ، $س_ن$ تمثل القيم العينية للمتغير العشوائى $س_ر$ المكون من ن وحدة، فالوسط الحسابى هو :

$$\bar{س} = \frac{س_١ + س_٢ + س_٣ + \cdots + س_ن}{ن} \qquad (١)$$

ويستخدم الدليل هنا (١ ، ٢ ، ٣، . . . ،ن) لبيان ترتيب اختيار العنصر، وهذا يعنى أن القيم العينية قد تكون غير مرتبة ترتيباً تصاعدياً أو تنازلياً. إذاً $س_١$ هى القيمة العينية الأولى و $س_ن$ هى القيمة العينية الأخيرة. وقد تكون $س_١$ مساوية $س_ن$ من حيث القيمة أو أكبر منها أو أصغر منها.

يمكن كتابة المعادلة السابقة بطريقة أفضل بعد إدخال رمز التجميع سيقما (Σ) وتعميم الرمز (ر) الذى لا يكون إلا عدداً صحيحاً يبدأ من الدليل الأول، وينتهى بالدليل الأخير (ن) لتصبح المعادلة على النحو الآتى :

$$\bar{س} = \frac{\sum_{ر=١}^{ن} س_ر}{ن} \qquad (٢)$$

وخلاصة ذلك هى أن الوسط الحسابى هو مجموع القيم العينية مقسوماً على عددها.

أما فى حالة التوزيعات التكرارية فتعتبر $س_ر$ هى مركز الفئة ذات التكرار $ك_ر$. ويكون الوسط الحسابى للبيانات المكونة من (ف) فئة هو :

$$\bar{س} = \frac{س_١ ك_١ + س_٢ ك_٢ + س_٣ ك_٣ + \cdots + س_ف ك_ف}{ك_١ + ك_٢ + ك_٣ + \cdots + ك_ف} \qquad (٣)$$

وبما أن $ك_١ + ك_٢ + ك_٣ + \cdots + ك_ف$ هى مجموع التكرارات الذى يساوى حجم العينة (ن)، فإن المعادلة الأخيرة يمكن كتابتها على النحو الآتى :

$$\bar{س} = \frac{\sum_{ر=١}^{ف} س_ر ك_ر}{ن} \qquad (٤)$$

حيث إن ف تعنى عدد الفئات.

تستخدم المعادلة (٢) لإيجاد الوسط الحسابى للمفردات بينما تستخدم المعادلة (٤) لإيجاد نفس الإحصائية فى حالة البيانات المبوبة. إذاً فناتج المعادلة (٤) هو تقريب للوسط الحسابى الصحيح الذى تعطيه المعادلة (٢)، بيد أن الفرق يكون ضئيلاً جداً، إذا تم التبويب بطريقة صحيحة. ويعزى ذلك الفرق إلى افتراض أن جميع التكرارات الخاصة بكل فئة واقعة فى منتصفها، أى أن عناصر كل فئة قد استبدلت بعنصر واحد هو منتصف الفئة.

مثال (١ , ٤) :

أوجد الوسط الحسابى للمفردات التالية :

١٦؛ ٢٢؛ ٢١؛ ٢٠؛ ٢٣؛ ٢١؛ ١٩؛ ١٥؛ ١٣؛ ٢٣؛ ١٧؛ ٢٠؛ ٢٩؛ ١٨؛ ٢٢؛ ١٦؛ ٢٥.

الحل :

ن = ١٧

$\sum$ س ر = ١٦ + ٢٢ + ٢١ + ٢٠ + ... + ٢٢ + ١٦ + ٢٥ = ٣٤٠

$\bar{س} = \frac{\sum س_ر}{ن}$

$= \frac{٣٤٠}{١٧}$

∴ $\bar{س}$ = ٢٠

البرنامج التالى يقوم بحساب الوسط الحسابى لمجموعة مفردات حسب المعادلة :

$$\bar{س} = \frac{\sum س_ر}{ن}$$

بمعنى أن :

$$M = \frac{S}{N}$$

عدد القيم هنا يرمز له بالرمز N

ومجموع المتغيرات بالرمز S

والوسط الحسابى بالرمز M

```
10  REM برنامج لحساب الوسط الحسابي لمجموعة مفردات
20  S=0
30  READ N , REM عدد القيم
35  PRINT ,'البيانات'
40  FOR I=1 TO N
50   READ X
60   PRINT ,X
70   S=S+X
80  NEXT I
90  PRINT
100 PRINT
110 M=S/N
120 PRINT ,M;'=الوسط الحسابي'
130 PRINT
140 PRINT
150 DATA 17,16,22,21,20,23,21,19,15,13,23,17,20,29,18,22,16,25
160 END
```

المخرجات

```
البيانات
16
22
21
20
23
21
19
15
13
23
17
20
29
18
22
16
25

20  =الوسط الحسابي
```

مثال (٢ ، ٤)

أوجد الوسط الحسابي للبيانات التالية :

رقم الفئة	الفئات العمرية (بالسنوات)	$ك_ر$	$س_ر$	$س_ر ك_ر$
١	١٥,٥-١٢,٥	٢	١٤	٢٨
٢	١٨,٥-١٥,٥	٣	١٧	٥١
٣	٢١,٥-١٨,٥	٤	٢٠	٨٠
٤	٢٤,٥-٢١,٥	٥	٢٣	١١٥
٥	٢٧,٥-٢٤,٥	١٤	٢٦	٣٦٤
٦	٣٠,٥-٢٧,٥	٢١	٢٩	٦٠٩
٧	٣٣,٥-٣٠,٥	٥٩	٣٢	١٨٨٨
٨	٣٦,٥-٣٣,٥	٢٤	٣٥	٨٤٠
٩	٣٩,٥-٣٦,٥	٧	٣٨	٢٦٦
١٠	٤٢,٥-٣٩,٥	٦	٤١	٢٤٦
١١	٤٥,٥-٤٢,٥	٢	٤٤	٨٨
١٢	٤٨,٥-٤٥,٥	١	٤٧	٤٧
١٣	٥١,٥-٤٨,٥	٢	٥٠	١٠٠
المجموع		١٥٠		٤٧٢٢

$$\bar{س} = \frac{\sum_{ر=1}^{١٣} س_ر ك_ر}{ن} \qquad (٤)$$

$$\sum_{ر=1}^{١٣} ك_ر = ن$$

$$\therefore ن = ١٥٠$$

$$\bar{س} = \frac{٤٧٢٢}{١٥٠}$$

$$\bar{س} = ٣١,٤٨ \text{ سنة}$$

البرنامج التالى يقوم بحساب الوسط الحسابى لبيانات مبوبة بطريقة مراكز الفئات، وباستخدام المعادلة :

$$M = \frac{D_1}{F_1}$$

حيث :

M = الوسط الحسابى ($\bar{س}$)

$D_1 = \sum_{I=1}^{N} C(I) \times F(I)$

$C(I)$ = مركز الفئة ($س_ر$)

$F(I)$ = تكرار الفئة ($ك_ر$)

$FI = \sum F(I)$ = ($\sum ك_ر$) مجموع التكرارات

N = عدد الفئات

```
10  REM  برنامج لحساب الوسط الحسابي لبيانات مجمعة
20  DIM A(13),B(13),C(13),D(13),E(13),F(13)
30  F1=0
40  READ N REM NO OF OBSERVATIONS
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)    REM الحد الادنى, الحد الاعلى والتكرار
70   F1=F1+F(I)             REM مجموع التكرارات
80   C(I)=(A(I)+B(I))/2
90   D(I)=C(I)*F(I)
100  D1=D1+D(I)
110 NEXT I
120 PRINT '      ناتج الضرب      مراكز الفئات   التكرار       المجموعات',
130 PRINT '
140 PRINT '      ________        ________       ________      __________  '
145 PRINT
150 FOR I=1 TO N
170  PRINT USING 190, D(I),C(I),F(I),B(I),A(I)
180 NEXT I
190 :    ####     ###     ##      ##.# - ##.#
200 PRINT '  ________  ________  ________  ________',
210 PRINT
220 PRINT TAB(6);D1;TAB(29);F1;TAB(40);'المجموع'
230 PRINT
240 PRINT
250 M=D1/F1
260 PRINT ,M;'=الوسط الحسابي'
270 PRINT
280 PRINT
290 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
300 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
310 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
320 END
```

المخرجات

ناتج الضرب	مراكز الفئات	التكرار	المجموعات
28	14	2	15.5 - 12.5
51	17	3	18.5 - 15.5
80	20	4	21.5 - 18.5
115	23	5	24.5 - 21.5
364	26	14	27.5 - 24.5
609	29	21	30.5 - 27.5
1888	32	59	33.5 - 30.5
840	35	24	36.5 - 33.5
266	38	7	39.5 - 36.5
246	41	6	42.5 - 39.5
88	44	2	45.5 - 42.5
47	47	1	48.5 - 45.5
100	50	2	51.5 - 48.5
4722		150	المجموع

31.48 =الوسط الحسابي

خواص الوسط الحسابي :

تهدف جميع مقاييس النزعة المركزية إلى تقدير نقطة (Point Estimation) معينة تمثل مركز توزيع البيانات. والوسط الحسابي هو أكثر هذه المقاييس استخداماً لما يتمتع به من مزايا استنتاجية كعدم التحيز في تقدير وسط المجتمع الذى سحبت منه العينة العشوائية، والكفاءة (Efficiency) ، إضافة إلى سهولة التعامل الجبرى. هذا، ولقد ساعدت سهولة عملياته

الجبرية، واستخدام رمز التجميع على أن تكون للوسط الحسابى ثلاث خواص هامة وهى :

أ ـ مجموع انحرافات القيم العينية عن وسطها الجسابى يساوى صفراً.

والانحراف فى الإحصاء يعنى الفرق أو البعد، فإذا كانت القيم هى

$\text{س}_1 ، \text{س}_2 ، \text{س}_3 ، \ldots\ldots ، \text{س}_{\text{ن}}$

فالانحرافات هى $\text{س}_1 - \bar{\text{س}}$ ، $\text{س}_2 - \bar{\text{س}}$ ، $\text{س}_3 - \bar{\text{س}}$ ، ...... ، $\text{س}_{\text{ن}} - \bar{\text{س}}$

وإذا كان مجموع هذه الانحرافات يساوى صفراً، فهذا يعنى أن :

$$\text{س}_1 - \bar{\text{س}} + \text{س}_2 - \bar{\text{س}} + \text{س}_3 - \bar{\text{س}} + \ldots + \text{س}_{\text{ن}} - \bar{\text{س}} = 0$$

وبإدخال رمز التجميع على الجانب الأيمن يصبح :

$$\sum_{\text{ر}=1}^{\text{ن}} (\text{س}_{\text{ر}} - \bar{\text{س}}) = 0 \qquad (5)$$

أى أن :

$$\sum \text{س}_{\text{ر}} - \sum \bar{\text{س}} = 0$$

$$\sum \text{س}_{\text{ر}} - \text{ن}\bar{\text{س}} = 0$$

$$\sum \text{س}_{\text{ر}} - \text{ن}\frac{\sum \text{س}_{\text{ر}}}{\text{ن}} = 0$$

$$\therefore \sum \text{س}_{\text{ر}} - \sum \text{س}_{\text{ر}} = 0 \qquad (6)$$

مثال (٣ , ٤) :

أوجد الوسط الحسابى للمتغيرات التالية ومجموع انحرافاتها عن وسطها الحسابى، والمتغيرات هى :

١٠، ١٢، ١٥، ١٣، ١٢، ١١، ١١

الحل :

$$\bar{س} = \frac{\sum س_{ر}}{ن} \qquad (٢)$$

$$\sum س_{ر} = ١٠ + ١٢ + ١٥ + ١٣ + ١٢ + ١١ + ١١$$

$$= ٨٤$$

$$ن = ٧$$

$$\bar{س} = \frac{٨٤}{٧}$$

$$\therefore \bar{س} = ١٢$$

	$س_{ر}$	$س_{ر} - \bar{س}$
	١٠	- ٢
	١٢	٠
	١٥	٣
	١٣	١
	١٢	٠
	١١	- ١
	١١	- ١
المجموع	٨٤	صفر

ب ـ مجموع مربعات الانحرافات للقيم العينية عن وسطها الحسابى أقل من مجموع مربعات الانحرافات لنفس القيم عن اى نقطة أخرى.

وهذه الخاصة هى التى جعلت الوسط الحسابى يكون أدق وأكفأ مقاييس النزعة المركزية .

ومجموع مربعات الانحرافات عن الوسط جبرياً يكون على النحو التالى :

$$\sum (س_{ر} - \bar{س})^{٢} \qquad (٧)$$

وهذه يمكن أن تكون :

$$\sum س_ر^{٢} - ٢ \bar{س} \sum س_ر + ن \bar{س}^{٢} \quad \ldots\ldots (٨)$$

لنأخذ قيمة فرضية أخرى تبعد عن $\bar{س}$ بمقدار ط. أى أن القيمة الفرضية هى $\bar{س} + ط$ وقد تكون قيمة ط موجبة أو سالبة. إذاً فمجموع مربعات الانحرافات عن القيمة الجديدة هى :

$$\sum (س_ر - \bar{س} - ط)^{٢}$$

$$= \sum (س_ر^{٢} + \bar{س}^{٢} + ط^{٢} - ٢ س_ر \bar{س} - ٢ س_ر ط + ٢ \bar{س} ط)$$

$$= \sum س_ر^{٢} + \sum \bar{س}^{٢} + \sum ط^{٢} - ٢ \bar{س} \sum س_ر - ٢ ط \sum (س_ر - \bar{س})$$

إلا أن

$$\sum (س_ر - \bar{س}) = ٠$$

وبالتالى :

$$\sum (س_ر - \bar{س} - ط)^{٢} = \sum س_ر^{٢} + ن \bar{س}^{٢} + ن ط^{٢} - ٢ \bar{س} \sum س_ر$$

وهذه تزيد على $\sum (س_ر - \bar{س})^{٢}$ المبينة فى (٨) سابقاً بزيادة $ن ط^{٢}$. وهذه القيمة موجبة سواء كانت ط موجبة أو سالبة.

مثال (٤ , ٤) :

استخدم البيانات الواردة فى مثال (٣)، وأوجد الفرق بين مجموع مربعات الانحرافات عن الوسط الحسابى، ومجموع مربعات الانحرافات عن قيمة فرضية أخرى مقدارها ١٠ وتحقق أن الفرق بين المجموعين يساوى $ن ط^{٢}$.

الحل :

س ر	س ر - سَ	س ر - ١٠	(س ر - سَ)٢	(س ر - ١٠)٢
١٠	-٢	٠	٤	٠
١٢	٠	٢	٠	٤
١٥	٣	٥	٩	٢٥
١٣	١	٣	١	٩
١٢	٠	٢	٠	٤
١١	-١	١	١	١
١١	-١	١	١	١
المجموع ٨٤	صفر	١٤	١٦	٤٤

سَ = ١٢

ن = ٠٧

∴ ط = ١٠ - ١٢

= -٢

ن ط٢ = ٧ × (-٢)٢

= ٢٨

Σ(س ر - ١٠)٢ - Σ(س ر - سَ)٢ = ٤٤ - ١٦

= ٢٨

∴ الفرق = ن ط٢ (٩)

جـ - يمكن إيجاد الوسط الحسابى من مجموع القيم العينية وعددها، دون الحاجة للتوزيع التكرارى أو المفردات.

بما أن

سَ = Σس ر / ن فى حالة المفردات

كما أن

$$\bar{س} = \frac{\sum س_ر ك_ر}{\sum ك_ر}$$ في حالة التوزيعات التكرارية

لذا أصبح من الممكن إيجاد قيمة $\bar{س}$ متى كانت $\sum س_ر$ و $ن$ أو $\sum س_ر ك_ر$ مع $\sum ك_ر$ معلومة. هذا، وسوف يلاحظ فيما بعد أن بقية مقاييس النزعة المركزية الهامة لا تمتلك هذه الصفة.

لهذه الصفة السهلة البسيطة عدة نتائج جانبية هامة تم اشتقاقها من المعادلة :

$$ن \bar{س} = \sum س_ر \qquad (١٠)$$

إذ أصبح بالإمكان معرفة مجموع القيم في حالات كثيرة، واستخدمت تلك المجاميع لإيجاد ما يسمى بالوسط الحسابي المرجح، ووسط مجاميع وفروق أزواج القيم المتناظرة.

$$\text{فالوسط الحسابي المرجح} = \frac{س_١ + س_٢ + س_٣ + \ldots + س_ف}{ك_١ + ك_٢ + ك_٣ + \ldots + ك_ف}$$

أي أن الوسط الحسابي المرجح هو

$$\frac{\sum_{ر=١}^{ف} س_ر ك_ر}{\sum ك_ر} \qquad (١١)$$

والذي يشبه الوسط الحسابي في حالة التوزيعات التكرارية، إلا أن $ك_ر$ تعني الوزن أو الأهمية. لذلك يمكن اعتبار الوسط الحسابي في حالة التوزيعات التكرارية، هو عبارة عن وسط حسابي مرجح بالتكرارات. وأكثر الحالات التي يستخدم فيها الوسط الحسابي المرجح هي إيجاد الوسط الحسابي للأسعار المرتبطة بكميات مختلفة عددها (ف).

مثال (٤,٥) :

تقوم إحدى المؤسسات ببيع ثلاثة أنواع من السيارات، فإذا كان سعر السيارة من النوع الأول (ماركة أ) يساوي ٢٥٠٠٠ ريال، والسعر من الماركة (ب) يساوي ٢٣٠٠٠ ريال، وسعر

السيارة من الماركة (جـ) يساوى ١٨٠٠٠ ريال، فأوجد متوسط سعر السيارة، علماً بأن الشركة قد باعت خلال العام الماضى عدد ١٠٠٠ و ١٤٠٠ و ١٦٠٠ سيارة من الماركة (أ) والماركة (ب) والماركة (جـ) على التوالى.

الوسط الحسابى غير المرجح هو $\frac{\sum \text{س}_{\text{ر}}}{\text{ن}}$

$$= \frac{١٨٠٠٠ + ٢٣٠٠٠ + ٢٥٠٠٠}{٣}$$

= ٢٢٠٠٠ ريال

إلا أن هذا الوسط غير صحيح لأن الكميات المبيعة تختلف من نوع إلى آخر، لذلك يجب استخدام الوسط الحسابى المرجح بالكميات على النحو التالى :

$$\text{الوسط الحسابى المرجح} = \frac{١٦٠٠ \times ١٨٠٠٠ + ١٤٠٠ \times ٢٣٠٠٠ + ١٠٠٠ \times ٢٥٠٠٠}{١٦٠٠ + ١٤٠٠ + ١٠٠٠}$$

$$= \frac{٨٦٠٠٠٠٠٠}{٤٠٠٠}$$

= ٢١٥٠٠ ريال

فيما يلى برنامج لحساب الوسط الحسابى المرجح للبيانات الواردة فى المثال السابق باستخدام المعادلتين :

$$A = \frac{T}{R} \quad ، \quad B = \frac{S}{3}$$

حيث :

$A =$ الوسط الحسابى غير المرجح

$S = \sum_{i=1}^{3} P_i$

$P =$ سعر السيارة

$T = \sum P_i F_i$

$F =$ عدد السيارات

$R = \sum F_i$

```
10  REM   برنامج لحساب الوسط الحسابي المرجح
20  READ P1,F1,P2,F2,P3,F3
30  PRINT ,'البيانات'
40  PRINT
50  PRINT ,P1
60  PRINT ,F1
70  PRINT ,P2
80  PRINT ,F2
90  PRINT ,P3
100 PRINT ,F3
110 S=P1+P2+P3
120 T=P1*F1+P2*F2+P3*F3
130 R=F1+F2+F3
140 A=S/3
150 B=T/R
160 PRINT
170 PRINT ,A;'= الوسط الحسابى غير المرجح'
190 PRINT
200 PRINT ,B;'= الوسط الحسابى المرجح'
210 PRINT
220 PRINT
230 DATA 25000,1000,23000,1400,18000,1600
240 END

                    المخرجات
                    --------
           البيانات

            25000
            1000
            23000
            1400
            18000
            1600

            22000   = الوسط الحسابى غير المرجح
            21500   = الوسط الحسابى المرجح
```

أما إذا توفرت لدينا الأوساط الحسابية لمجموعات جزئية، وأردنا الحصول على الوسط الحسابى الكلى لهذه المجموعات، أى الوسط الحسابى للمجموعة التى تضم كل هذه الأجزاء، فإنه يساوى :

$$\bar{\text{س}} = \frac{\bar{\text{س}}_{١}\,\text{ك}_{١} + \bar{\text{س}}_{٢}\,\text{ك}_{٢} + \bar{\text{س}}_{٣}\,\text{ك}_{٣} + \ldots\ldots + \bar{\text{س}}_{\text{ف}}\,\text{ك}_{\text{ف}}}{\text{ك}_{١} + \text{ك}_{٢} + \text{ك}_{٣} + \ldots\ldots + \text{ك}_{\text{ف}}}$$

حيث ف هى عدد الأجزاء.

$$\therefore \bar{\text{س}} = \frac{\sum_{\text{ر}=١}^{\text{ف}} \text{س}_{\text{ر}}\,\text{ك}_{\text{ر}}}{\sum \text{ك}_{\text{ر}}} \qquad (١٢)$$

حيث $\bar{\text{س}}_{\text{ر}}$ هو الوسط الحسابى للمجموعة ر، $\text{ك}_{\text{ر}}$ هو عدد وحدات (متغيرات) المجموعة ر.

مثال (٦,٤) :

كان الوسط الحسابى لمادة الرياضيات فى ثلاثة فصول هو ٨٠؛ ٨٦؛ ٧٢ فأوجد الوسط الحسابى الكلى إذا كان عدد الطلاب فى الفصول الثلاثة على التوالى يساوى ٣٠ و ٣٥ و ٦٠ طالباً.

$$\bar{س} = \frac{\sum \bar{س}_ر ك_ر}{\sum ك_ر} \qquad (١٢)$$

$$= \frac{٨٠ \times ٣٠ + ٨٦ \times ٣٥ + ٧٢ \times ٦٠}{٣٠ + ٣٥ + ٦٠}$$

$$\bar{س} = \frac{٩٧٣٠}{١٢٥}$$

$$\therefore \bar{س} = ٧٧,٨٤ \text{ درجة}$$

أما إذا كانت هناك مجموعة من القيم المتناظرة، وكانت كل مجموعة مستقلة عن الأخرى وعدد مفرداتها مساوياً لها، فالوسط الحسابى لمجموع الظاهرتين ($\bar{س}$) هو :

$$\bar{س} = \bar{س}_١ + \bar{س}_٢ \qquad (١٣)$$

كما أن الوسط الحسابى للفرق بين الظاهرتين هو :

$$\bar{س} = \bar{س}_١ - \bar{س}_٢ \qquad (١٤)$$

حيث $\bar{س}_١$، $\bar{س}_٢$ هما الوسطان الحسابيان للظاهرتين.

٣ - الوسيط (و ٥٠٪) :

إذا رتبت القيم العينية ترتيباً تصاعدياً على النحو التالى :

$س_١ \leqslant س_٢ \leqslant س_٣ \leqslant \ldots \leqslant س_ن$ أو ترتيباً تنازليا، فالإحصائية :

١) $س_١$ تسمى بالقيمة الصغرى.

٢) $\text{س}_{\text{ن}}$ تسمى بالقيمة الكبرى.

٣) $\text{س}_{\text{ن}} - \text{س}_{1}$ تسمى المدى.

٤) $\text{س}_{\frac{(\text{ن}+1)}{2}}$ تسمى الوسيط إذا كان عدد المتغيرات فردياً.

٥) $\frac{1}{2}\left(\text{س}_{\frac{\text{ن}}{2}} + \text{س}_{\frac{(\text{ن}+2)}{2}}\right)$ تسمى الوسيط إذا كان عدد المتغيرات زوجياً.

إذاً الوسيط هو القيمة الوسطى للمقادير المرتبة ترتيباً تصاعدياً أو تنازلياً، كما أنه الوسط الحسابى للقيمتين الوسطيين إذا كان عدد المتغيرات زوجياً.

البرنامج التالى يقوم بحساب الوسيط للمفردات، والبرنامج يحتوى على فقرة لفرز القيم تنازلياً (السطر 45 إلى السطر 110).

```
10  REM   برنامج لحساب الوسيط لمجموعة مفردات
20  DIM X(17)
25  READ N   REM  عدد القيم
30  FOR I=1 TO N
35   READ X(I)
40  NEXT I
45  REM    فرز القيم تنازليا
50  FOR I=1 TO N-1
60   FOR J=I+1 TO N
65    IF X(I)>X(J) THEN 100
70    T=X(I)
80    X(I)=X(J)
90    X(J)=T
100  NEXT J ---
110 NEXT I
120 PRINT ,'الترتيب','القيمة'
130 PRINT '________';'________'
140 FOR I=1 TO N
150  PRINT ,I,X(I)
160 NEXT I
170 PRINT
180 P=INT((N+1)/2)
190 M=X(P)
200 PRINT 'الوسيط هو المتغير رقم';P;' وهو الرقم';M
210 DATA 17,13,25,16,36,17,18,19,16,22,19,38,41,21,36,81,17,44
220 END
```

المخرجات

الترتيب	القيمة
1	81
2	44
3	41
4	38
5	36
6	36
7	25
8	22
9	21
10	19
11	19
12	18
13	17
14	17
15	16
16	16
17	13

الوسيط هو المتغير رقم 9 وهو الرقم 21

مثال (٤,٧) :

أوجد الوسيط للمتغيرات :

١٦، ٢٢، ٢١، ٢٠، ٢٣، ٢١، ١٩، ١٥، ١٣، ٢٣، ١٧، ٢٠، ٢٩، ١٨، ٢٢،
١٦، ٢٥

الحل :

١) ترتب القيم ترتيباً تصاعدياً (أو تنازلياً) لتصبح على النحو التالى :

١٣، ١٥، ١٦، ١٦، ١٧، ١٨، ١٩، ٢٠، ٢٠، ٢١، ٢١، ٢٢، ٢٢، ٢٣، ٢٣،
٢٥، ٢٩

٢) بما أن عدد المفردات فردى (ن = ١٧)، لذا فترتيب الوسيط هو :

$$\frac{١٧ + ١}{٢} = ٩$$

فالوسيط هو المتغير التاسع .

٣) و٥٠٪ = ٢٠

أما البيانات المبوبة فهى مرتبة ترتيباً تصاعدياً حسب الفئات، ويجب تحديد الفئة التى تحتوى على الوسيط (الفئة الوسيطية) أولاً وقبل كل شيء . والفئة الوسيطية هى تلك الفئة التى يعلو عندها التجمع التكرارى الصاعد لنصف مجموع التكرارات ($\frac{ن}{٢}$) لأول مرة، وبافتراض أن :

ح أ   تعنى الحد الأدنى للفئة الوسيطية .
ط   طول الفئة الوسيطية .
ك   تكرار الفئة الوسيطية .
ن´   التجمع التكرارى الصاعد لدى الفئة التى تسبق الفئة الوسيطية مباشرة

فإن :

$$و_{٥٠٪} = ح_{أ} + \frac{\left(\frac{ن}{٢} - ن'\right) ط}{ك} \qquad (١٥)$$

مثال (٤,٨) :

أوجد الوسيط للبيانات الواردة فى مثال (٢) والمبينة فيما يلى :

رقم الفئة	الفئات العمرية	$ك_ر$	التجمع التكرارى الصاعد
١	١٢,٥ - ١٥,٥	٢	٢
٢	١٥,٥ - ١٨,٥	٣	٥
٣	١٨,٥ - ٢١,٥	٤	٩
٤	٢١,٥ - ٢٤,٥	٥	١٤
٥	٢٤,٥ - ٢٧,٥	١٤	٢٨
٦	٢٧,٥ - ٣٠,٥	٢١	٤٩
٧	٣٠,٥ - ٣٣,٥	٥٩	١٠٨
٨	٣٣,٥ - ٣٦,٥	٢٤	١٣٢
٩	٣٦,٥ - ٣٩,٥	٧	١٣٩
١٠	٣٩,٥ - ٤٢,٥	٦	١٤٥
١١	٤٢,٥ - ٤٥,٥	٢	١٤٧
١٢	٤٥,٥ - ٤٨,٥	١	١٤٨
١٣	٤٨,٥ - ٥١,٥	٢	١٥٠
المجموع		١٥٠	

$$\frac{\text{ن}}{٢} = \frac{١٥٠}{٢}$$

$$= ٧٥$$

∴ الفئة الوسيطية هى الفئة السابعة = ٣٠,٥ - ٣٣,٥

وعليه تكون :

$\text{ح}_1 = ٣٠,٥$ ؛ $\bar{\text{ن}} = ٤٩$ ؛ $\text{ط} = ٣$ ؛ $\text{ك} = ٩٥$

$$\text{و}_{٥٠\%} = \text{ح}_1 + \frac{\left(\frac{\text{ن}}{٢} - \bar{\text{ن}}\right)\text{ط}_1}{\text{ك}}$$

$$= \frac{٣ \times (٤٩ - ٧٥)}{٥٩} + ٣٠,٥$$

$$\therefore و_{٥٠\%} = ٣١,٨٢٢$$

البرنامج التالى يقوم بحساب الوسيط لبيانات مبوبة باستخدام طريقة التجمع التكرارى الصاعد، والتى سبق شرحها وباستخدام المعادلة :

$$R = L + \frac{(Y-O)P}{Q}$$

حيث :

R = الوسيط

L = الحد الأدنى للفئة الوسيطية

$Y = \frac{T}{2}$

T = مجموع التكرارات (ن)

P = طول الفئة الوسيطية= B – A

B = الحد الأعلى للفئة الوسيطية

A = الحد الأدنى للفئة الوسيطية

Q = تكرار الفئة الوسيطية

O = التجمع التكرارى الصاعد السابق للوسيطية

```
10  REM  برنامج لحساب الوسيط لبيانات مجمعه
20  DIM A(13),B(13),C(13),D(13),E(13),F(13),G(13)
30  T=0
40  READ N     REM عدد القيم
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)    REM  التكرار , الحد الاعلى , الحد الادنى
70   T=T+F(I)
72  NEXT I
73  C(1)=F(1)
75  FOR I=2 TO N
80   C(I)=C(I-1)+F(I)    REM   التجمع التكرارى الصاعد
110 NEXT I
120 PRINT '         التجمع التكرارى       التكرار      الفئه       رقم
130 PRINT '           الصاعد                                       الفئه'
140 PRINT '        ____________      _______     _______      _____'
145 PRINT
150 FOR I=1 TO N
170  PRINT USING 190, C(I),F(I),B(I),A(I),I
180 NEXT I
190 :             ###              ##      ##.# - ##.#  ##
200 PRINT '
210 PRINT
```

```
220 PRINT TAB(33);T; TAB(60); 'المجموع'
230 PRINT
240 PRINT
250 Y=T/2
260 FOR I=1 TO N
270  IF C(I)>Y THEN 290
280  J=I
290 NEXT I
300 PRINT B(J+1);'-';A(J+1);'وهى';J+1;'الفئه الوسيطيه هى الفئه رقم'
310 L=A(J+1)
320 O=C(J)
330 P=B(J)-A(J)
340 Q=F(J+1)
350 R=L+((Y-O)*P)/Q
360 PRINT
370 PRINT ,R;'= الوسيط'
380 PRINT
390 PRINT
400 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
410 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
420 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
430 END
```

المخرجات

رقم الفئه	الفئه	التكرار	التجمع التكرارى الصاعد
1	15.5 - 12.5	2	2
2	18.5 - 15.5	3	5
3	21.5 - 18.5	4	9
4	24.5 - 21.5	5	14
5	27.5 - 24.5	14	28
6	30.5 - 27.5	21	49
7	33.5 - 30.5	59	108
8	36.5 - 33.5	24	132
9	39.5 - 36.5	7	139
10	42.5 - 39.5	6	145
11	45.5 - 42.5	2	147
12	48.5 - 45.5	1	148
13	51.5 - 48.5	2	150
المجموع		150	

الفئه الوسيطيه هى الفئه رقم 7 وهى 30.5 - 33.5

الوسيط = 31.82202

٤ ـ خصائص الوسيط واستخداماته :

١ ـ سهل التعريف وسهل الحساب، ذلك لأنه لا يعتمد على القيم العينية، وإنما يستخدم الرتب لهذه القيم . وبالرغم من أن الرتب تتزايد مع القيم العينية إلا أن الوسيط هو الأفضل لتحديد المرتبة الوسطى .

٢ ـ لايتأثر بالقيم المتطرفة (الشاذة)، والقيمة الشاذة هى التى تختلف اختلافاً كبيراً عن القيمة الصغرى أو الكبرى التى تليها بعد ترتيب المقادير، أى إنها كبيرة جداً أو صغيرة

جداً، مقارنة ببقية القيم. خذ على سبيل المثال البيانات الواردة فى المثال (٦) وافرض أن القيمة الأخيرة كانت ٢٥٠ وليست ٢٥ .

فالوسيط لا يزال كماهو = ٢٠

أما الوسط الحسابى فقد أصبح

$$\bar{س} = \frac{\sum س_ر}{ن} \qquad (٢)$$

$$= \frac{٥٦٥}{١٧}$$

$$\therefore \bar{س} = ٣٣,٢٤$$

أى أن الوسط قد أصبح أعلى من ١٦ قيمة من بين ١٧ قيمة، ذلك لأن قيمة الوسط الحسابى تتغير إذا أضيفت أى متغيرات تختلف فى قيمها عن الوسط الحسابى السابق للبيانات قبل الإضافة. لذلك يستخدم الوسيط لوصف الأجور فى مؤسسات معينة أو الدخل؛ لأن القيم المتطرفة تكون من أصل البيانات فى هذه الحالة، واستخدام الوسيط يضمن أن نصف الأفراد قد حصلوا على أجور أقل من قيمته.

٣ - يمكن استخراج الوسيط فى حالة الفئات المفتوحة أو المعلومات الناقصة التى يعرف ترتيبها، ذلك لأن الوسيط لا يحتاج لمراكز الفئات ولا القيم العينية ذاتها ما دام ترتيبها معلوماً. ولعل هذه الخاصية هى أهم خصائص الوسيط التى جعلت استخداماته ضرورية فى بعض المجالات، كالأعمار التى لايمكن تحديدها، والدخل، ودرجات الحرارة فى بعض الحالات. كذلك يستخدم الوسيط كثيراً فى المجال الصناعى، خاصة فى الفحوصات التى تحتاج لإتلاف بعض القطع، إلا أنه يفترض انتظام التوزيع داخل الفئات، بينما يفترض الوسط الحسابى أن التوزيع داخل الفئات طبيعى.

٥ - المنوال (ل):

هو القيمة الأكثر تكراراً، والمنوال التقريبى للبيانات المبوبة هو مركز الفئة التى يقابلها أكبر تكرار، لذلك تسمى هذه الفئة بالفئة المنوالية.

إذاً قد يكون هناك أكثر من منوال فى حالة المفردات التى تتساوى من حيث عدد التكرارات، وقد لا يوجد منوال إطلاقاً إن لم تكن هناك تكرارات.

مثال (٤,٩) :

أوجد المنوال فى كل من الحالات التالية :

أ ـ ١٤ ، ١٦ ، ١٦ ، ١٨ ، ١٩ ، ١٩ ، ١٩ ، ١٢ ، ١٠ ، ٢٤ ، ٢٦

ب ـ ١٤ ، ١٦ ، ١٦ ، ١٦ ، ١٨ ، ١٨ ، ١٨ ، ١٩ ، ١٩ ، ١٩ ، ١٢ ، ١٠ ، ٢٤ ، ٢٦ .

جـ ـ ١٤ ، ١٦ ، ١٨ ، ١٩ ، ١٢ ، ١٠ ، ٢٤ ، ٢٦

الحل :

أ / المنوال = ١٩ لأنها الأكثر تكراراً.

ب / لها ثلاثة منوالات وهى ١٦ ، ١٨ ، ١٩

جـ / ليس لها منوال.

أما البيانات المبوبة فقد تكون لها أيضاً أكثر من فئة منوالية ، وتعتبر طريقة بيرسون لحساب المنوال فى كل حالة هى الأكثر استخداماً وتشبه إلى حد كبير طريقة استخراج الوسيط. فإذا كانت :

ف_1 تعنى الفرق بين تكرار الفئة المنوالية والتكرار لدى الفئة التى تسبقها، أى أن $\text{ف}_1 = \text{ك} - \text{كَ}$ حيث كَ هى تكرار الفئة السابقة للفئة المنوالية.

ف_2 تعنى ك - كُ حيث كُ هى تكرار الفئة اللاحقة للفئة المنوالية.

$\text{ح}_{\text{أ}}$ تعنى الحد الأدنى للفئة المنوالية.

ط تعنى طول الفئة المنوالية. فالمعادلة هى :

$$\text{ل} = \text{ح}_{\text{أ}} + \left(\frac{\text{ف}_1}{\text{ف}_1 + \text{ف}_2}\right)\text{ط} \qquad (16)$$

مثال (٤,١٠) :

أوجد المنوال للبيانات الواردة فى المثال (٢) :

(أ) المنوال التقريبى هو مركز الفئة المنوالية.

الفئة المنوالية ٣٠,٥ - ٣٣,٥

∴ المنوال التقريبى = ٣٢

(ب) باستخدام المعادلة :

$$\text{ل} = \text{ح}_{\text{أ}} + \left(\frac{\text{ف}_1}{\text{ف}_1 + \text{ف}_2}\right) \text{ط}$$

$\text{ف}_1 = ٥٩ - ٢١ = ٣٨$ ؛ $\text{ف}_2 = ٣٥$

$$\therefore \text{ل} = ٣٠,٥ + \left(\frac{٣٨}{٣٨ + ٣٥}\right) \times ٣$$

$$\therefore \text{ل} = ٣٢,٠٦٢$$

فيما يلى برنامج لحساب المنوال لبيانات تكرارية وباستخدام المعادلة

$$M = L + \left(\frac{F_1}{F_1 + F_2}\right) C$$

حيث :

M = المنوال

L = الحد الأدنى للفئة المنوالية

F_1 = الفرق بين تكرار الفئة المنوالية والسابقة عليها

F_2 = الفرق بين تكرار الفئة المنوالية واللاحقة لها

```
10  REM برنامج لحساب المنوال لبيانات مجمعه
20  DIM A(13),B(13),F(13)
30  T=0       REM  مجموع التكرارات
40  READ N    REM  عدد القيم
50  PRINT USING 70
60  PRINT USING 80
70  :      التكرار           الفئـــه
80  :      _______           ________
85  PRINT
90  FOR I=1 TO N
100  READ A(I),B(I),F(I)   REM  الحد الادنى , الحد الاعلى , التكرار
110  T=T+F(I)
115  PRINT USING 130, F(I),B(I),A(I)
120 NEXT I
130 :          ##          ##.# - ##.#
140 PRINT
150 PRINT USING 170
160 PRINT USING 180,T
170 :          ___
180 :          ###  المجموع
190 REM   لايجاد التكرار الاعلى
200 H=F(1)
```

```
210 FOR I=2 TO N
220  IF F(I)<H THEN 250
230  H=F(I)
240  J=I
250 NEXT I
260 F1=F(J)-F(J-1)
270 F2=F(J)-F(J+1)
280 L=A(J)
290 C=B(J)-A(J)
300 M=L+F1/(F1+F2)*C
310 PRINT
320 PRINT USING 325,B(J),A(J)
325 ;      ##.# - ##.# الفئه المنواليه هى الفئه
330 PRINT
340 PRINT ,M;'= المنوال'
350 PRINT
360 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
370 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
380 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
390 END
```

المخرجات

الفئــــه	التكرار
15.5 - 12.5	2
18.5 - 15.5	3
21.5 - 18.5	4
24.5 - 21.5	5
27.5 - 24.5	14
30.5 - 27.5	21
33.5 - 30.5	59
36.5 - 33.5	24
39.5 - 36.5	7
42.5 - 39.5	6
45.5 - 42.5	2
48.5 - 45.5	1
51.5 - 48.5	2
المجموع	150

الفئه المنواليه هى الفئه 30.5 - 33.5

المنوال = 32.06163

٦ ـ خصائص المنوال واستخداماته :

تعتبر استخدامات المنوال قليلة، مقارنة بالمقياسين السابقين، وذلك لعدة أسباب، منها أنه قد لايوجد منوال للبيانات القليلة، كما أنه قد يوجد أكثر من منوال بين البيانات المبوبة، بسبب الاختلاف فى اختيار الفئات لنفس البيانات، إضافة إلى وجود أكثر من منوال لنفس التوزيع التكرارى فى بعض الحالات. إلا أنه يتميز على الوسيط والوسط بأنه :

١ ـ يمكن استخدامه لاستخراج مقياس أفضل للنزعة المركزية إذا كانت البيانات وصفية.

٢ ـ بالرغم من عدم قابليته للعمليات الجبرية إلا أنه يسهل تقريبه، سواء بطريقة بيرسون أو بمركز الفئة المنوالية، لذلك قد يستخدم كمقياس تقريبى للنزعة المركزية.

٣ـ لا يتأثر بالقيم المتطرفة، ويمكن استخراجه فى حالة الفئات أو المجموعات المفتوحة التى يندر أن تكون فئات منوالية. أضف إلى ذلك أنه قد يكون أفضل من الوسط الحسابى إذا كان توزيع البيانات بعيداً عن التماثل (شديد الالتواء)؛ لأن القيم الشاذة تؤثر كثيراً على الوسط الحسابى فى هذه الحالة.

٧ـ العلاقة بين الوسط والوسيط والمنوال :

إذا كان التوزيع متماثلاً تماماً، كما لو كان توزيع البيانات يمثل انعكاساً من المرآة، فالمقاييس الثلاثة متساوية، وتساوى جميعها مركز الفئة المنوالية.

أما إذا كان التوزيع التكرارى ملتوياً التواء بسيطاً، وله فئة منوالية (قمة تكرارية) واحدة، فيكون الوسيط بين المنوال والوسط الحسابى. هذا ولقد أثبت بيرسون العلاقة التقريبية التالية بتكرار التجارب على توزيعات ليست شديدة الالتواء. والعلاقة هى :

$$\frac{(\text{الوسط الحسابى} - \text{المنوال})}{٣} = \text{الوسط الحسابى} - \text{الوسيط}$$

أى أن :

$$\bar{س} - ل \quad = \quad ٣\ (\bar{س} - و_{٥٠\%}) \qquad (١٧)$$

مثال (٤,١١)

أوجد المنوال باستخدام الوسط الحسابى والوسيط للبيانات الواردة فى المثال (٢)، وقارن ذلك بالمنوال المستخرج فى المثال (٩).

$$\bar{س} - ل = ٣\ (\bar{س} - و_{٥٠\%}) \qquad (١٧)$$

$$\therefore ل = \bar{س} - ٣\ (\bar{س} - و_{٥٠\%})$$

$$ل = ٣١,٤٨ - ٣\ (٣١,٤٨ - ٣١,٨٢٢)$$

$$ل = ٣٢,٥٠٦$$

وبما أن هذه القيمة تختلف عن تلك التى استخرجت فى المثال (٩)، والتى كانت تساوى ٣٢,٠٦٢ فهذا يعنى أن الالتواء ليس بسيطاً جداً.

وتجدر الإشارة هنا إلى أن الوسط الحسابى يكون أصغر المقاييس الثلاثة، إذا كان الالتواء سالباً، وأكبرها إذا كان موجباً، بينما يقع الوسيط بين الوسط الحسابى والمنوال.

٨ ـ الوسط الهندسى (هـ) :

هو الجذر النونى لحاصل ضرب ن قيمة عينية. أى أن الوسط الهندسى (هـ) للمتغيرات $س_١$، $س_٢$، $س_٣$، . . . . ،$س_ن$ هو :

$$هـ = \sqrt[ن]{س_١ \times س_٢ \times س_٣ \times . . . . \times س_ن} \qquad (١٨)$$

وهذا يعنى أن :

$$لو هـ = \frac{لو س_١ + لو س_٢ + لو س_٣ + ... + لو س_ن}{ن} \qquad (١٩)$$

$$لو هـ = \frac{\sum_{ر=١}^{ن} لو س_ر}{ن} \qquad (٢٠)$$

إذاً فلوغاريتم الوسط الهندسى للقيم هو الوسط الحسابى للوغاريتمات تلك القيم.

مثال (١٢، ٤) :

أوجد الوسط الهندسى للقيم :

١٦، ٢٢، ٢١، ٢٠، ٢٣، ٢١، ١٩، ١٥، ١٣، ٢٣، ١٧، ٢٠، ٢٩، ١٨، ٢٢، ١٦، ٢٥

$$لو هـ = \frac{لو ١٦ + لو ٢٢ + لو ٢١ + ... + لو ٢٥}{١٧}$$

$$لو هـ = \frac{١,٢٠٤١٢ + ١,٣٤٢ + ١,٣٢٢ + ... + ١,٣٩٨}{١٧}$$

$$لو هـ = \frac{٢١,٩٧٧٨٨٧}{١٧}$$

$$لو هـ = ١,٢٩٢٨١٦٨$$

$$\therefore هـ = ١٩,٦٢٥$$

البرنامج أدناه يقوم بحساب الوسط الهندسى لمفردات، باستخدام المعادلة :

$$G = \sqrt[N]{T}$$

حيث :

G = الوسط الهندسى

N = عدد المتغيرات

$$T = T_1 \times T_2 \times T_3 \times \ldots\ldots \times T_N$$

```
10  REM   برنامج لحساب الوسط الهندسى لمجموعه مفردات
20  T=1
25  PRINT ,'البيانات'
27  PRINT ,'________'
30  READ N  REM   عدد الارقام
40  FOR I=1 TO N
50   READ X
60   PRINT ,X
70   T=T*X
80  NEXT I
90  G=T**(1/N)
100 PRINT
110 PRINT
120 PRINT ,G;'= الوسط الهندسى'
130 PRINT
140 PRINT
150 DATA 17,16,22,21,20,23,21,19,15,13,23,17,20,29,18,22,16,25
160 END
```

المخرجات

```
البيانات
________
16
22
21
20
23
21
19
15
13
23
17
20
29
18
22
16
25

19.62531   = الوسط الهندسى
```

أما فى حالة التوزيعات التكرارية حيث $س_ر$ تعنى مركز الفئة ر التى يساوى تكرارها $ك_ر$.

وبما أن $\sum ك_ر = ن$ فالوسط الهندسى (هـ) هو :

$$هـ = \sqrt[ن]{س_١^{ك_١} \times س_٢^{ك_٢} \times س_٣^{ك_٣} \times \ldots س_ر^{ك_ر} \times \ldots\ldots \times س_ف^{ك_ف}} \qquad (٢١)$$

حيث ف تعنى عدد المجموعات .

$$\therefore لو هـ = \frac{ك_١ لو س_١ + ك_٢ لو س_٢ + ك_٣ لو س_٣ + \ldots + ك_ر لو س_ر + \ldots + ك_ف لو س_ف}{ن} \qquad (٢٢)$$

$$لو هـ = \frac{\sum_{ر=١}^{ف} ك_ر لو س_ر}{ن} \qquad (٢٣)$$

مثال (٤,١٣) :

أوجد الوسط الهندسى للبيانات الواردة فى المثال (٢) والمبينة أدناه :

رقم الفئة	الفئة	$ك_ر$	$س_ر$	لو١٠ $س_ر$	$ك_ر$ لو $س_ر$
١	١٥,٥-١٢,٥	٢	١٤	١,١٤٦	٢,٢٩٢
٢	١٨,٥-١٥,٥	٣	١٧	١,٢٣٠	٣,٦٩٠
٣	٢١,٥-١٨,٥	٤	٢٠	١,٣٠١	٥,٢٠٤
٤	٢٤,٥-٢١,٥	٥	٢٣	١,٣٦٢	٦,٨١٠
٥	٢٧,٥-٢٤,٥	١٤	٢٦	١,٤١٥	١٩,٨١٠
٦	٣٠,٥-٢٧,٥	٢١	٢٩	١,٤٦٢	٣٠,٧٠٢
٧	٣٣,٥-٣٠,٥	٥٩	٣٢	١,٥٠٥	٨٨,٧٩٥
٨	٣٦,٥-٣٣,٥	٢٤	٣٥	١,٥٤٤	٣٧,٠٥٦
٩	٣٩,٥-٣٦,٥	٧	٣٨	١,٥٨٠	١١,٠٦٠
١٠	٤٢,٥-٣٩,٥	٦	٤١	١,٦١٣	٩,٦٧٨
١١	٤٥,٥-٤٢,٥	٢	٤٤	١,٦٤٣	٣,٢٨٦
١٢	٤٨,٥-٤٥,٥	١	٤٧	١,٦٧٢	١,٦٧٢
١٣	٥١,٥-٤٨,٥	٢	٥٠	١,٦٩٩	٣,٣٩٨
	المجموع	١٥٠			٢٢٣,٤٥٣

$$\text{لو هـ} = \frac{\sum_{ر=1}^{13} ك_ر \, لو س_ر}{ن} \qquad (23)$$

$$\text{لو هـ} = \frac{223,453}{150}$$

$$\text{لو هـ} = 1,4896866$$

$$\therefore \text{هـ} = 30,881$$

البرنامج التالى يقوم بحساب الوسط الهندسى لبيانات تكرارية :

```
10  REM  برنامج لحساب الوسط الهندسي لبيانات مجمعه
20  DIM A(13),B(13),C(13),D(13),E(13),F(13),G(13)
30  F1=0      REM  مجموع التكرارات
35  D1=0      REM  D مجموع
40  READ N     REM عدد المجموعات
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)   REM  الحد الادنى , الحد الاعلى , التكرار
70   F1=F1+F(I)
80   C(I)=(A(I)+B(I))/2
90   G(I)=LGT(C(I))
100  D(I)=G(I)*F(I)
110  D1=D1+D(I)
120 NEXT I
125 PRINT ' ____________________________________________________________'
130 PRINT '  لوسر كرلوسر     لوسر          سـر        كر          الفئـــه     رقم'
140 PRINT '                                                              الفئه'
150 PRINT ' ____________________________________________________________'
155 PRINT
160 FOR I=1 TO N
180  PRINT USING 190, D(I),G(I),C(I),F(I),B(I),A(I),I
190 ' ##.###    #.###        ###         ##        ##.# - ##.#  ##
200 NEXT I
210 PRINT ' ______                        ___   '
220 PRINT
230 PRINT D1; TAB(33);F1; TAB(50); 'المجموع'
240 PRINT
250 PRINT
260 H=D1/F1
270 M=10**H
310 PRINT ,M;'=الوسط الهندسى'
320 PRINT
330 PRINT
340 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
350 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
360 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
370 END
```

المخرجات

رقم الفئة	الفئة	ك ر	س ر	لو س ر	ك ر لو س ر
1	15.5 - 12.5	2	14	1.146	2.292
2	18.5 - 15.5	3	17	1.230	3.691
3	21.5 - 18.5	4	20	1.301	5.204
4	24.5 - 21.5	5	23	1.362	6.809
5	27.5 - 24.5	14	26	1.415	19.810
6	30.5 - 27.5	21	29	1.462	30.710
7	33.5 - 30.5	59	32	1.505	88.804
8	36.5 - 33.5	24	35	1.544	37.058
9	39.5 - 36.5	7	38	1.580	11.058
10	42.5 - 39.5	6	41	1.613	9.677
11	45.5 - 42.5	2	44	1.643	3.287
12	48.5 - 45.5	1	47	1.672	1.672
13	51.5 - 48.5	2	50	1.699	3.398
المجموع		150			223.4698

الوسط الهندسى= 30.88858

٩ ـ خصائص الوسط الهندسى واستخداماته :

الخاصية الأساسية للوسط الهندسى هى أنه عبارة عن قيمة تحويلية (Transformed) للوسط الحسابى. هذا ولقد تم اشتقاقه ليتم تطبيقه فى حالة معينة بدلاً من الوسط الحسابى ، وتلك الحالة هى التى تتبع فيها البيانات نمط المتوالية الهندسية التزايدية أو التناقصية. والمتوالية الهندسية هى مجموعة من القيم المرتبة بحيث تكون النسبة بين كل قيمتين متتاليتين كمية ثابتة ، وبالتالى يمكن الانتقال فيها من أى قيمة س ر إلى القيمة التالية س ر + ١ بالضرب فى الكمية الثابتة. فمثلا القيم التالية عبارة عن متوالية هندسية كميتها الثابتة تساوى ٣.

٢ ، ٦ ، ١٨ ، ٥٤، ١٦٢

فالوسط الحسابى لهذه المتوالية = ٤٨,٤

وهو يبدو وكأن ١٦٢ قيمة شاذة مع أنها جزء من المتوالية، أما الوسط الهندسى لنفس البيانات أعلاه فهو يساوى ١٨ ، وهى فعلاً القيمة التى تتوسط هذه القيم.

بذلك يصبح الوسط الهندسى هو المقياس الأفضل للنزعة المركزية فى حالات الزيادة أو النقصان بنسب ثابتة، كما هو الحال فى تقديرات التعداد السكانى والأسعار.

وباختصار :

الوسط الهندسى هو الأفضل فى جميع الحالات التى يمكن أن تستخدم فيها قاعدة الفائدة المركبة ؛ لإيجاد الجملة (جـ) التى يؤول إليها مبلغ من المال (أ) بعد (ن) فترة زمنية بمعدل فائدة ع٪ عن كل فترة على النحو التالى :

$$جـ = أ(١ + ع)^{ن} \qquad (٢٤)$$

أضف إلى ذلك أن الوسط الهندسى هو الأفضل، لإيجاد متوسط التغير النسبى عند استخدام الأرقام القياسية.

١٠ ـ الوسط التوافقى (ق) :

إذا كانت $س_١$، $س_٢$، $س_٣$، . . . . ، $س_ن$ هى قيم عينية، فالوسط التوافقى هو :

$$ق = \frac{ن}{\frac{١}{س_١} + \frac{١}{س_٢} + \frac{١}{س_٣} + \cdots + \frac{١}{س_ن}} \qquad (٢٥)$$

$$ق = \frac{ن}{\sum (\frac{١}{س_ر})} \qquad (٢٦)$$

فهو إذاً عدد المتغيرات مقسوما على مجموع مقلوبات المتغيرات.

مثال (١٤ , ٤) :

أوجد الوسط التوافقى للمتغيرات :

١٦، ٢٢، ٢١، ٢٠، ٢٣، ٢١، ١٩، ١٥، ١٣، ٢٣، ١٧، ٢٠، ٢٩، ١٨، ٢٢، ١٦، ٢٥

الحل :

$$ق = \frac{١٧}{\frac{١}{١٦} + \frac{١}{٢٢} + \frac{١}{٢١} + \cdots + \frac{١}{٢٥}}$$

$$\text{ق} = \frac{\text{١٧}}{\text{٠,٠٤} + \ldots + \text{٠,٠٤٧٦} + \text{٠,٠٤٥٥} + \text{٠,٠٦٢٥}}$$

$$\text{ق} = \frac{\text{١٧}}{\text{٠,٨٨٣١٨٦٢}}$$

$$\text{ق} = \text{١٩,٢٤٨}$$

فيما يلى برنامج لحساب الوسط التوافقى لقيم عينية باستخدام المعادلة :

$$M = \frac{N}{T}$$

حيث :

الوسط التوافقى = M

عدد المتغيرات = N

مجموع مقلوبات المتغيرات = T

```
10  REM  برنامج لحساب الوسط التوافقى لمجموعه مفردات
20  T=0     REM  مجموع مقلوب البيانات
30  PRINT ,'البيانات'
40  PRINT ,'__________'
50  READ N  REM    عدد الارقام
60  FOR I=1 TO N
70   READ X
80   PRINT ,X
90   T=T+1/X
100 NEXT I
110 PRINT
120 PRINT
130 M=N/T
140 PRINT ,M;'= الوسط التوافقى'
150 PRINT
160 PRINT
170 DATA 17,16,22,21,20,23,21,19,15,13,23,17,20,29,18,22,16,25
180 END
```

المخرجات

البيانات
16
22
21
20
23
21
19
15
13
23
17
20
29
18
22
16
25

الوسط التوافقى = 19.24847

أما فى حالة البيانات المبوبة فى جدول تكرارى مكون من (ف) فئة فالوسط التوافقى هو :

$$\text{ق} = \frac{\text{ك}_{1} + \text{ك}_{2} + \text{ك}_{3} \cdots\cdots\cdots + \text{ك}_{\text{ف}}}{\frac{\text{ك}_{1}}{\text{س}_{1}} + \frac{\text{ك}_{2}}{\text{س}_{2}} + \frac{\text{ك}_{3}}{\text{س}_{3}} + \cdots\cdots\cdots + \frac{\text{ك}_{\text{ف}}}{\text{س}_{\text{ف}}}} \qquad (٢٧)$$

$$\text{ق} = \frac{\sum_{\text{ر}=1}^{\text{ف}} \text{ك}_{\text{ر}}}{\sum \left(\frac{\text{ك}_{\text{ر}}}{\text{س}_{\text{ر}}} \right)} \qquad (٢٨)$$

$$\therefore \text{ق} = \frac{\text{ن}}{\sum \left(\frac{\text{ك}_{\text{ر}}}{\text{س}_{\text{ر}}} \right)} \qquad (٢٩)$$

مثال (٤,١٥) :

أوجد الوسط التوافقى للبيانات الواردة فى المثال (٢) والمبينة أدناه :

رقم الفئة	الفئة	$ك_ر$	$س_ر$	$\frac{ك_ر}{س_ر}$
١	١٢,٥-١٥,٥	٢	١٤	٠,١٤٣
٢	١٥,٥-١٨,٥	٣	١٧	٠,١٧٦
٣	١٨,٥-٢١,٥	٤	٢٠	٠,٢٠٠
٤	٢١,٥-٢٤,٥	٥	٢٣	٠,٢١٧
٥	٢٤,٥-٢٧,٥	١٤	٢٦	٠,٥٣٨
٦	٢٧,٥-٣٠,٥	٢١	٢٩	٠,٧٢٤
٧	٣٠,٥-٣٣,٥	٥٩	٣٢	١,٨٤٤
٨	٣٣,٥-٣٦,٥	٢٤	٣٥	٠,٦٨٦
٩	٣٦,٥-٣٩,٥	٧	٣٨	٠,١٨٤
١٠	٣٩,٥-٤٢,٥	٦	٤١	٠,١٤٦
١١	٤٢,٥-٤٥,٥	٢	٤٤	٠,٠٤٥
١٢	٤٥,٥-٤٨,٥	١	٤٧	٠,٠٢١
١٣	٤٨,٥-٥١,٥	٢	٥٠	٠,٠٤٠
	المجموع	١٥٠		٤,٩٦٦

$$ق = \frac{١٥٠}{٤,٩٦٦}$$

$$\therefore ق = ٣٠,٢٠٥$$

أما فى حالة التوزيعات التكرارية فالبرنامج التالى يقوم بحساب الوسط التوافقى، وتستخدم كمثال البيانات الواردة فى المثال السابق وباستخدام المعادلة :

$$M = \frac{FI}{DI}$$

حيث :

M = الوسط التوافقى

FI = مجموع التكرارات (ن)

DI = مجموع مناسيب التكرارات لمراكز الفئات

```
10   REM  برنامج لحساب الوسط التوافقي لبيانات مجمعة
20   DIM A(13),B(13),C(13),D(13),F(13)
30   F1=0        REM  مجموع التكرارات
35   D1=0        REM  D مجموع
40   READ N      REM  عدد المجموعات
50   FOR I=1 TO N
60    READ A(I),B(I),F(I)    REM  الحد الادنى , الحد الاعلى , التكرار
70    F1=F1+F(I)
80    C(I)=(A(I)+B(I))/2
90    D(I)=F(I)/C(I)
100   D1=D1+D(I)
120  NEXT I
125  PRINT '         ________________________________________________________'
130  PRINT '         كر/سر          سر          كر          الفئـــه       رقم
140  PRINT '                                                              الفئة'
150  PRINT '         ________________________________________________________'
155  PRINT
160  FOR I=1 TO N
170   PRINT USING 180, D(I),C(I),F(I),B(I),A(I),I
180  :          ##.###         ###         ##         ##.# - ##.#  ##
190  NEXT I
200  PRINT '        ________              ___        '
210  PRINT
220  PRINT USING 230,D1,F1
230  :          ##.###                    ####               المجموع
240  PRINT
250  PRINT
300  M=F1/D1
310  PRINT ,M;'=الوسط التوافقي'
320  PRINT
330  PRINT
340  DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
350  DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
360  DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
370  END
```

المخرجات

كر/سر	سر	كر	الفئـــه	رقم الفئة
0.143	14	2	15.5 - 12.5	1
0.176	17	3	18.5 - 15.5	2
0.200	20	4	21.5 - 18.5	3
0.217	23	5	24.5 - 21.5	4
0.538	26	14	27.5 - 24.5	5
0.724	29	21	30.5 - 27.5	6
1.844	32	59	33.5 - 30.5	7
0.686	35	24	36.5 - 33.5	8
0.184	38	7	39.5 - 36.5	9
0.146	41	6	42.5 - 39.5	10
0.045	44	2	45.5 - 42.5	11
0.021	47	1	48.5 - 45.5	12
0.040	50	2	51.5 - 48.5	13
4.966		150	المجموع	

الوسط التوافقي= 30.205

١١ ـ خصائص الوسط التوافقى واستخداماته :

إذا كانت $س_١$ ، $س_٢$ ، $س_٣$ ، .... ، $س_ن$ هى قيم عينية فمقلوبات هذه القيم هى :

$$\frac{١}{س_١} ، \frac{١}{س_٢} ، \frac{١}{س_٣} ، .... ، \frac{١}{س_ن}$$

أما مجموع هذه المقلوبات فهو :

$$\frac{١}{س_١} + \frac{١}{س_٢} + \frac{١}{س_٣} + ... + \frac{١}{س_ن}$$

$$= \sum_{ر=١}^{ن} \frac{١}{س_ر}$$

وأما الوسط الحسابى لمقلوبات هذه القيم ، فهو مجموعها مقسوماً على عددها (ن) . إذاً هو :

$$\bar{س} = \frac{\left(\sum \frac{١}{س_ر} \right)}{ن} \qquad (٣٠)$$

ومقلوب الوسط الحسابى $\left(\frac{١}{\bar{س}} \right)$ لمقلوبات هذه القيم هو :

$$\frac{١}{\bar{س}} = \frac{ن}{\left(\sum \frac{١}{س_ر} \right)} \qquad (٣١)$$

$$\therefore \frac{١}{\bar{س}} = ق \qquad (٣٢)$$

إذاً فالوسط التوافقى هو مقلوب الوسط الحسابى لمقلوبات القيم.وعليه ، فالوسط التوافقى هو قيمة تحويلية للوسط الحسابى يستخدم بدلاً منه فى حالات خاصة جداً ، شأنه فى ذلك شأن الوسط الهندسى .

هناك بعض الحالات التى تكون فيها القيم عبارة عن ناتج قسمة متغير على متغير آخر ليس من نفس وحدة القياس. فالسرعة هى ناتج قسمة المسافة على الزمن (كم/ الساعة)، والسعر هو ناتج قسمة المبلغ على عدد القطع مثلاً (ريال/ قطعة)، والإنتاجية هى ناتج قسمة الإنتاج على المساحة (طن/ هكتار).

فالزمن المتوسط لقطع مسافة كيلو متر واحد، أو متوسط عدد القطع التى يمكن شراؤها بريال واحد، أو متوسط المساحة التى يجب زراعتها لإنتاج طن واحد، يعنى تحويل البسط إلى وحدة واحدة. فتكون وحدات القياس السالفة الذكر على النحو التالى :

$$\frac{١}{\text{ساعة}} ، \frac{١}{\text{قطعة}} ، \frac{١}{\text{هكتار}}$$

وكل واحدة منها تمثل مقلوباً لقيمة معينة. ففى مثل هذه الحالات وما شابهها يكون الوسط التوافقى هو المقياس الأفضل بدلاً من الوسط الحسابى.

١٢ ـ الربيعات والعشيرات والمئينيات :

(Quartiles, Deciles and percentiles)

الربيعات هى التى تقسم القيم إلى أربعة أقسام يساوى كل منها الربع (٢٥٪)، لذلك فهى ثلاثة :

أ ـ الربيع الأعلى (و ٧٥٪) :

وهو الذى يقسم القيم إلى جزأين بحيث يكون عدد القيم التى أقل منه يساوى $\frac{٣}{٤}$ والربع الباقى أكثر منه. ويمكن استخراجه من البيانات المبوبة حسب القاعدة :

$$(\text{و}_{٧٥٪}) = \text{ح}_{\text{أ}} + \frac{(\frac{٣}{٤}\text{ن} - \text{ن}^{-})}{\text{ك}}\,\text{ط} \qquad (٣٣)$$

حيث :

$\text{ح}_{\text{أ}}$ هى فئة الربيع الأعلى التى يعلو عندها التجمع التكرارى الصاعد لقيمة $\frac{٣}{٤}$ ن لأول مرة.

ن´ هى التجمع التكرارى الصاعد لدى الفئة التى تسبق فئة الربيع الأعلى.

ط هى طول فئة الربيع الأعلى.

ك هى تكرار فئة الربيع الأعلى.

ويلاحظ أن القاعدة نفسها هى قاعدة الوسيط مع اختلاف تفسير الرموز.

ب - الربيع الأوسط (و ٥٠٪) :

وهو الوسيط.

جـ - الربيع الأدنى (و ٢٥٪) :

وهو القيمة التى تعلو $\frac{١}{٤}$ القيم بينما تعلوعليها $\frac{٣}{٤}$ تلك القيم، وبذلك تكون قاعدة الربيع الأدنى هى :

$$\text{و}_{٢٥\%} = \text{ح}_{١} + \frac{\left(\frac{\text{ن}}{٤} - \text{ن}'\right)\text{ط}}{\text{ك}} \qquad (٣٤)$$

حيث

ح١ هى الحد الأدنى لفئة الربيع الأدنى التى يعلوعندها التجمع التكرارى الصاعد لقيمة $\frac{\text{ن}}{٤}$ لأول مرة.

ن´ هى التجمع التكرارى الصاعد لدى الفئة التى تسبق فئة الربيع الأدنى.

ط هى طول فئة الربيع الأدنى.

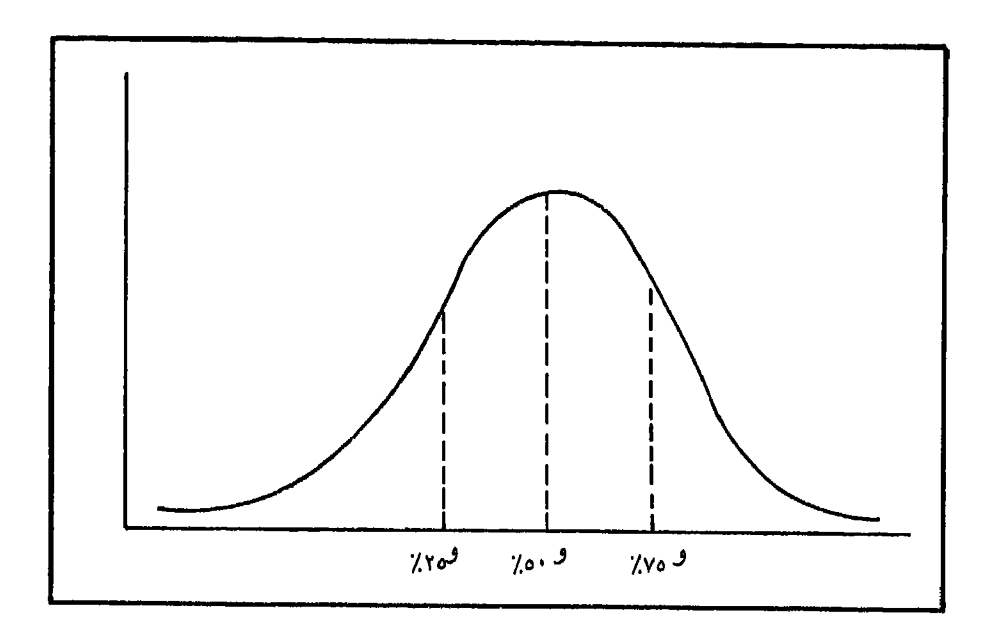
ك تكرار فئة الربيع الأدنى.

وهذا يعنى أن :

$$\text{و}_{٢٥\%} < \text{و}_{٥٠\%} < \text{و}_{٧٥\%} \qquad (٣٥)$$

أما و٧٥٪ – و٥٠٪ فهى لاتساوى و٥٠٪ – و٢٥٪ إلا إذا كان التوزيع الخاص بالبيانات متماثلاً تماماً.

إذا كانت الربيعات هى التى تقسم المساحة التى تقع تحت المضلع التكرارى إلى أربعة أقسام متساوية.



فإن العشيرات هى التى تقسم تلك المساحة إلى عشرة أقسام متساوية، والمئينيات هى التى تقسمها إلى مائة قسم متساوٍ. فالمئينى الأول هو و١٪ والعشير الأول هو و١٠٪. ويكون المئينى الخمسون هو العشير الخامس، وهو الربيع الثانى، وهو الوسيط، لأنها جميعاً تساوى و٥٠٪. ولاستخراج الجزىء الرائىّ تستخدم المعادلة :

$$\text{و}_{\text{ر}}\% = \text{ح}_{1} + \frac{\left(\frac{\text{ر ن}}{100} - \text{ن}'\right)\text{ط}}{\text{ك}} \qquad (٣٦)$$

وتجدر الإشارة هنا إلى أن هذه المقاييس ـ باستثناء الوسيط ـ ليست من مقاييس النزعة المركزية، ولكنها تستخدم لوصف التوزيع التكرارى للبيانات، وتحديد المواقع النسبية للمفردات مقارنة ببقية عناصر المجموعة. لذلك تستخدم كثيراً لتحديد التقديرات الخاصة بالطلاب.

مثال (١٦ , ٤) :

استخدم البيانات الواردة فى مثال (٧) والمبينة بعد لاستخراج الربيع الأعلى والربيع الأدنى .

رقم الفئة	الفئات العمرية	ك ر	التجمع الصاعد
١	١٢,٥ - ١٥,٥	٢	٢
٢	١٥,٥ - ١٨,٥	٣	٥
٣	١٨,٥ - ٢١,٥	٤	٩
٤	٢١,٥ - ٢٤,٥	٥	١٤
٥	٢٤,٥ - ٢٧,٥	١٤	٢٨
٦	٢٧,٥ - ٣٠,٥	٢١	٤٩ ← فئة الأدنى
٧	٣٠,٥ - ٣٣,٥	٥٩	١٠٨ ← فئة الوسيط
٨	٣٣,٥ - ٣٦,٥	٢٤	١٣٢ ← فئة الأعلى
٩	٣٦,٥ - ٣٩,٥	٧	١٣٩
١٠	٣٩,٥ - ٤٢,٥	٦	١٤٥
١١	٤٢,٥ - ٤٥,٥	٢	١٤٧
١٢	٤٥,٥ - ٤٨,٥	١	١٤٨
١٣	٤٨,٥ - ٥١,٥	٢	١٥٠
المجموع		١٥٠	

و ٧٥٪ :

$$\frac{٣}{٤}\,\text{ن} = \frac{١٥٠ \times ٣}{٤} = ١١٢,٥$$

$$\therefore \text{و}\,٧٥٪ = ٣٣,٥ + \frac{(١١٢,٥ - ١٠٨) \times ٣}{٢٤}$$

$$\therefore \text{و}\,٧٥٪ = ٣٤,٠٦٣$$

و ٢٥٪ :

$$\frac{\text{ن}}{٤} = ٣٧,٥$$

$$\therefore \text{و}\,٢٥٪ = ٢٧,٥ + \frac{(٣٧,٥ - ٢٨) \times ٣}{٢١}$$

$$\therefore \text{و}\,٢٥٪ = ٢٨,٨٥٧$$

البرنامج التالى يقوم باستخراج الآتى مستخدماً البيانات بالمثال السابق :

- فئة الربيع الأعلى .
- الفئة الوسيطية .
- فئة الربيع الأدنى .
- الربيع الأعلى .
- الوسيط .
- الربيع الأدنى .

باستخدام المعادلة العامة :

$$Q(I) = L(I) + \frac{(S(I) - O(I))\,P(I)}{W(I)}$$

وتكون القيمة هى الربيع الأعلى أو الوسيط أو الربيع الأدنى، عندما تكون I تساوى ١ أو ٢ أو ٣ على التوالى.

حيث :

Q = الربيع

L = الحد الأدنى للفئة

S = النسبة من عدد المتغيرات

D = المتجمع التكرارى للفئة السابقة

P = طول الفئة

W = تكرار الفئة

```
10  REM   برنامج لحساب الربيعات لبيانات مجمعة
20  DIM A(13),B(13),C(13),D(13),E(13),F(13),G(13)
30  T=0
40  READ N
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)
70   T=T+F(I)
80  NEXT I
90  C(1)=F(1)
100 FOR I=2 TO N
110  C(I)=C(I-1)+F(I)
120 NEXT I
125 PRINT '          ________________________________________________________'
130 PRINT '          التجمع التكرارى      التكرار      الفئـــة      رقم'
140 PRINT '            المتصاعد                                    الفئة'
150 PRINT '          ________________________________________________________'
160 PRINT
170 FOR I=1 TO N
180  PRINT USING 190, C(I),F(I),B(I),A(I),I
190 :              ###            ##        ##.# - ##.#  ##
200 NEXT I
210 PRINT '                                  ___     '
220 PRINT
230 PRINT TAB(33);T; TAB(50); 'المجموع'
240 PRINT
250 PRINT
260 S1=T*3/4
270 S2=T/2
280 S3=T/4
290 FOR I=1 TO N
300  IF C(I)>S1 THEN 320
310  J=I
320 NEXT I
330 FOR I=1 TO N
340  IF C(I)>S2 THEN 360
350  K=I
360 NEXT I
370 FOR I=1 TO N
380  IF C(I)>S3 THEN 400
390  V=I
400 NEXT I
410 PRINT USING 470, B(J+1),A(J+1),J+1
420 PRINT
430 PRINT USING 480, B(K+1),A(K+1),K+1
440 PRINT
450 PRINT USING 490, B(V+1),A(V+1),V+1
460 PRINT
470 :        ##.# - ##.#  وهى  ##  فئه الربيع الاعلى هى الفئه رقم
480 :        ##.# - ##.#  وهى  ##  الفئة الوسيطيه هى الفئه رقم
490 :        ##.# - ##.#  وهى  ##  فئه الربيع الادنى هى الفئه رقم
500 GOSUB 630
510 Q1=FNA(L1,S1,O1,P1,W1)
520 PRINT ,Q1; '= الربيع الاعلى'
525 PRINT
530 GOSUB 680
540 Q2=FNA(L2,S2,O2,P2,W2)
550 PRINT ,Q2; '= الوسيط'
560 PRINT
570 GOSUB 730
580 Q3=FNA(L3,S3,O3,P3,W3)
590 PRINT ,Q3; '= الربيع الادنى'
600 PRINT
610 PRINT
620 STOP
630 L1=A(J+1)
640 O1=C(J)
650 P1=B(J)-A(J)
660 W1=F(J+1)
670 RETURN
680 L2=A(K+1)
690 O2=C(K)
700 P2=B(K)-A(K)
710 W2=F(K+1)
720 RETURN
730 L3=A(V+1)
740 O3=C(V)
750 P3=B(V)-A(V)
760 W3=F(V+1)
770 RETURN
780 DEF FNA(L,X,O,P,W)=L+((X-O)*P)/W
810 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
820 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
830 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
840 END
```

المخرجات

رقم الفئه	الفئــه	التكرار	التجمع التكرارى الصاعد
1	15.5 - 12.5	2	2
2	18.5 - 15.5	3	5
3	21.5 - 18.5	4	9
4	24.5 - 21.5	5	14
5	27.5 - 24.5	14	28
6	30.5 - 27.5	21	49
7	33.5 - 30.5	59	108
8	36.5 - 33.5	24	132
9	39.5 - 36.5	7	139
10	42.5 - 39.5	6	145
11	45.5 - 42.5	2	147
12	48.5 - 45.5	1	148
13	51.5 - 48.5	2	150
المجموع		150	

فئه الربيع الاعلى هى الفئه رقم 8 وهى 36.5 - 33.5

الفئه الوسيطه هى الفئه رقم 7 وهى 33.5 - 30.5

فئه الربيع الادنى هى الفئه رقم 6 وهى 30.5 - 27.5

الربيع الاعلى = 34.0625

الوسيط = 31.82202

الربيع الادنى = 28.85713

تماريـن

١ ـ البيانات التالية تمثل عينات لرواتب عدد من العاملين فى أربع إدارات مختلفة بإحدى المؤسسات (بالريال) :

الرقم	الإدارة (أ)	الإدارة (ب)	الإدارة (جـ)	الإدارة (د)
١	٣٠٠٠	٦٣١٤	٤٩٥٠	٨٤٠٠
٢	٧٠٠٠	٨٤٠٧	١٢٣٦٠	١١٩٠٠
٣	٥١٠٠	٧١١٩	٩٠٩٠	٦٣٠٠
٤	٩٤٥٠	٦٤٣٣	١٠٧١٠	٤٩٠٠
٥	٨٧٦٥	١٠٥٢٨	٢٣٤٠	٣٥٠٠
٦	٦٤٠٠	٩٣١٧	٩٤٥٠	٧٠٠٠
٧	١٢٠٠٠	٧٧١٤	٦٣٩٠	٤٢٠٠
٨	٤٦٥٠		٦١٢٠	
٩	٦٦٣٥		٨١٩٠	
١٠	٨٣٠٠			

فأوجد الوسط الحسابى لكل إدارة.

٢ ـ استخدم بيانات السؤال الأول لإيجاد الوسط الحسابى لجميع أفراد العينات البالغ عددهم ٣٣ شخصاً.

٣ ـ أوجد وسط الأوساط الأربعة لبيانات السؤال الأول، وبين مدى اختلافه عن الوسط الخاص بالسؤال الثانى، ووضح سبب الفرق بين الوسطين.

٤ ـ أوجد الوسط الحسابى الخاص بالإدارتين (ب) و (د) معاً، مستخدماً بيانات السؤال الأول.

٥ ـ أوجد الوسط الحسابى للإدارتين (أ) و (جـ) معاً، مستخدماً بيانات السؤال الأول.

٦ ـ أوجد حجم عينة من الأعمار مجموع متغيراتها ١٦١ عاماً ووسطها الحسابى ٢٣ سنة.

٧ ـ الوسط الحسابى لعدد أيام انشغال السرير فى أحد الأجنحة ٩ أيام؛ بينما كان الوسط الحسابى لعدد أيام انشغال السرير فى جناح آخر ١٤ يوماً. أوجد الوسط الحسابى للجناحين معاً، إذا علمت أن العينة التى سحبت من الجناح الأول ٢٥ مريضاً، بينما كان قوام حجم العينة فى الجناح الآخر ٢١ مريضاً.

٨ ـ أوجد الوسط الحسابى للبيانات التالية الخاصة بتوزيع بعض المصابين فى حوادث المرور حسب الأعمار، والبيانات هى :

العمر بالسنوات	عدد المصابين
١ ـ ٦	١
٧ ـ ١٢	٧
١٣ ـ ١٨	١٥
١٩ ـ ٢١	٢٧
٢١ ـ ٢٥	٢١
٢٥ ـ ٣٠	٩
٣١ ـ ٣٨	٥
٣٩ ـ ٤٨	٢
٤٩ ـ ٥٨	١

٩ ـ أوجد الوسيط لكل مجموعة من المجموعات الواردة فى السؤال الأول.

١٠ ـ أوجد الوسيط للإدارات الأربع الواردة فى السؤال الأول، ووضح سبب اختلافه عن الوسيط للأربعة وسيطات.

١١ ـ أوجد الوسيط للإدارتين (ب) و (د) معاً مستخدماً بيانات السؤال الأول.

١٢ ـ أوجد الوسيط للإدارتين (أ) و (جـ) معاً مستخدماً بيانات السؤال الأول.

١٣ ـ قارن قيمة الوسيط فى كل من الأسئلة ٩ ـ ١٢ بنظيرتها الخاصة بالوسط الحسابى.

١٤ ـ أوجد الوسيط للبيانات الواردة فى السؤال الثامن. هل تختلف قيمة الوسيط عن الوسط الحسابى لتلك البيانات؟ ولماذا؟

١٥ ـ ماهى مزايا الوسيط على الوسط الحسابى؟

١٦ ـ أوجد المنوال لكل مجموعة من المجموعات الأربع الواردة فى السؤال الأول (إن وجد).

١٧ ـ أوجد المنوال للمجموعات الأربع الواردة فى السؤال الأول.

١٨ ـ أوجد المنوال للبيانات الواردة فى السؤال الثامن، وحدد اتجاه التواء تلك البيانات.

١٩ ـ استخدم بيانات السؤال الثامن لإيجاد ما يلى :

أ ـ الربيع الأدنى.
ب ـ الربيع الأعلى.
جـ ـ العشير الأعلى.
د ـ المئينى الأعلى.
هـ ـ العشير الأدنى.
و ـ السديس الأدنى.

٢٠ ـ كان سعر كيلو اللحم البقرى فى عواصم دول الخليج فى نفس الشهر على النحو الآتى :
٤,٥ دولار فى المدينة (أ).
٦,٠ دولارات فى المدينة (ب).
٣,٥ دولار فى المدينة (جـ).
٥,٠ دولارات فى المدينة (د).
٦,٥ دولار فى المدينة (هـ).
٨ دولارات فى المدينة (و).
أوجد الوسط المناسب للسعر بين تلك المدن.

٢١ ـ مجموع مربعات ٩ قيم عينية عن وسطها الحسابى يساوى ٦٤. فإذا كان الوسط الحسابى يساوى ١٢ فأوجد :
أ ـ مجموع الانحرافات عن الوسط الحسابى.
ب ـ مجموع الانحرافات عن قيمة أخرى مقدارها ٨.
جـ ـ مجموع مربعات الانحرافات عن قيمة أخرى مقدارها ١١.
د ـ مجموع الانحرافات عن قيمة مقدارها ١٦.
هـ ـ مجموع مربعات الانحرافات عن قيمة مقدارها ١٤.

٢٢ ـ بلغ تعداد السكان فى إحدى المدن ٦٠٠٠٠ شخص خلال عام ١٤٠٠ هـ. وباعتبار أن معدل الزيادة السنوية ٥٪، فإن التعداد للسنوات الأربع التالية يكون على النحو الآتى :
سكان ١٤٠١ هـ = ٦٣٠٠٠ نسمة.
سكان ١٤٠٢ هـ = ٦٦١٥٠ نسمة.
سكان ١٤٠٣ هـ = ٦٩٤٥٨ نسمة.
سكان ١٤٠٤ هـ = ٧٢٩٣٠ نسمة.
أوجد الوسط المناسب لعدد السكان خلال الخمس سنوات.

٢٣ ـ اكتب برامج بيسك لإيجاد حلول الأسئلة من (١) إلى (٥).

٢٤ ـ اكتب برنامج بيسك لإيجاد الوسط الحسابى للبيانات الواردة فى السؤال (٨).

٢٥ ـ اكتب برنامج بيسك لإيجاد الوسيط لكل مجموعة من المجموعات الواردة فى السؤال (١).

٢٦ ـ اكتب برنامج بيسك لإيجاد المنوال للمجموعات الأربع الواردة فى السؤال (١).

٢٧ ـ باستخدام البيانات الواردة فى السؤال (٨) اكتب برنامج بيسك لإيجاد الآتى :
١ ـ الربيع الأدنى. ٢ ـ الربيع الأعلى.

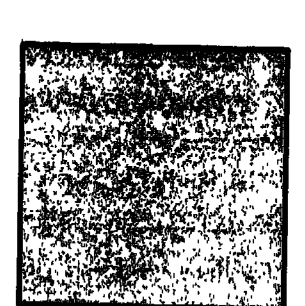


الفصل الخامس

مقاييس التشتت والعزوم

(Measures of Dispersion and Moments)





مقاييس التشتت والعزوم
(Measures of Dispersion and Moments)

١ ـ المقدمة :

يستخدم مقياس النزعة المركزية لتحديد قيمة نموذجية تتمركز حولها بقية القيم، إلا أن ذلك ليس كافياً لتوضيح الوصف الخاص بالتوزيع التكرارى للبيانات، أو مقارنتها بأى بيانات أخرى. خذ على سبيل المثال المجموعتين التاليتين من البيانات الفرضية :

س ر	ص ر
٣	١
٤	٤
٥	٧

فالوسط الحسابى للمجموعة الأولى يساوى الوسط الحسابى للمجموعة الثانية = ٤ .

غير أن المجموعتين مختلفتان تماماً؛ فالواضح أن مدى المجموعة ص أكبر من مدى المجموعة س. فالفرق بين قيمة الوسط الحسابى والقيم الأخرى فى المجموعة ص يعادل ثلاثة أمثال الفرق المناظر له فى المجموعة س. إذاً فتغيرات قيم المجموعة ص حول وسطها، أكبر من تغيرات قيم المجموعة س حول وسطها. أى أن المجموعة الثانية أكثر تباعداً، أو تبايناً أو تشتتاً من المجموعة الأولى.

إذاً لابد من مقاييس كمية لمدى تشتت البيانات فيما بينها، أو حول أى نقطة أخرى؛ لأن مقاييس النزعة المركزية لم توجد أساساً لتعطى تقديرات خاصة بتجانس القيم أو تشتتها.

هذا وتسمى مجموعة الإحصائيات الخاصة بالمقاييس الكمية للتشتت بمقاييس التشتت، وأهم هذه المقاييس هى :

١ ـ المدى RANGE

٢ ـ الانحراف الربيعى QUARTILE DEVIATION

٣ - الانحراف المتوسط MEAN DEVIATION
٤ - الانحراف المعيارى STANDARD DEVIATION

٢ - المدى :

ورد فى تعريف الوسيط أنه إذا رتبت القيم ترتيباً تصاعدياً بحيث إن :

$$س_١ \leqslant س_٢ \leqslant س_٣ \leqslant \dots \leqslant س_ن$$

فإن الإحصائية :

أ - $س_١$ تسمى القيمة الصغرى.
ب - $س_ن$ تسمى القيمة الكبرى.
جـ - $س_ن - س_١$ تسمى المدى. (١)

إذاً فالمدى هو الفرق بين أكبر قيمة وأصغر قيمة، إلا أنه بالرغم من سهولته لايستخدم إلا نادراً؛ لأنـه مقياس تقـريبى لايأخـذ فى الاعتبار إلا قيمتين فقط قد تكونان متطرفتين، كما لايمكن استخراجه فى حالة الفئات المفتوحة. وربما يعتبر أكثر الحالات التى يستخدم فيها المدى هى تكوين الجداول التكرارية، وضبط جودة الإنتاج فى المجال الصناعى (خرائط المراقبة).

وقـد يكـون المـدى هو الفـرق بين الحد الأدنى للفئة الأولى، والحد الأعلى للفئة العليا (الأخيرة)، فى حالة البيانات المبوبة، وقد يكون الفرق أيضاً بين مركز الفئة الأخيرة، ومركز الفئة الأولى، وهناك بعض الحالات التى يستبعد فيها جزء من البيانات ويستخرج المدى لبقية الأجـزاء. فإذا استبعـدت أعـلى وأدنى ١٠٪ من البيانات كان المدى هو الفرق بين المثينى التسعين والمثينى العاشر وسمى المدى المثينى. وهذه تستخدم كثيراً لاستبعاد الحالات المتطرفة كما هو الحال فى تقديرات الطلاب. أما إذا استبعد الربع الأعلى والربع الأدنى من البيانات فالمدى هنا هو الفرق بين الربيع الثالث ($و_{٧٥٪}$) والربيع الأول ($و_{٢٥٪}$) وهو ما يسمى بالمدى الربيعى. والمدى المحسوب بعد استبعاد أى قيمة، أو نسبة من البيانات، هو أحد شبيهات المدى.

٣ - الانحراف الربيعى :

هو نصف المدى الربيعى، وبذلك يكون

$$\text{الانحراف الربيعى} = \frac{(و_{٧٥٪} - و_{٢٥٪})}{٢} \qquad (٢)$$

وعليه، فالانحراف الربيعي يعتمد على الجزء الأوسط من ٥٠٪ من البيانات، فإذا كان العيب الرئيسى للمـدى هو الاعتماد الكلى على قيمتين متطرفتين أحياناً، فالعيب الرئيسى للانحراف الربيعى هو الإهمال التام لجزء من القيم، فاستبدل عدم التأثر بالقيم الشاذة بعدم دقة المقياس مقارنة بالمقاييس التالية. ويعتبر الانحراف الربيعى مفيداً جداً فى حالة التوزيعات ذات الفئات المفتوحة، والتى يمكن استخراج انحرافاتها الربيعية دون المقاييس الأخرى.

مثال (١,٥) :

أوجد المدى والانحراف الربيعى للبيانات المبوبة التالية :

رقم الفئة	الفئات	ك ر	س ر	التجمع الصاعد
١	١٥,٥-١٢,٥	٢	١٤	٢
٢	١٨,٥-١٥,٥	٣	١٧	٥
٣	٢١,٥-١٨,٥	٤	٢٠	٩
٤	٢٤,٥-٢١,٥	٥	٢٣	١٤
٥	٢٧,٥-٢٤,٥	١٤	٢٦	٢٨
٦	٣٠,٥-٢٧,٥	٢١	٢٩	٤٩
٧	٣٣,٥-٣٠,٥	٥٩	٣٢	١٠٨
٨	٣٦,٥-٣٣,٥	٢٤	٣٥	١٣٢
٩	٣٩,٥-٣٦,٥	٧	٣٨	١٣٩
١٠	٤٢,٥-٣٩,٥	٦	٤١	١٤٥
١١	٤٥,٥-٤٢,٥	٢	٤٤	١٤٧
١٢	٤٨,٥-٤٥,٥	١	٤٧	١٤٨
١٣	٥١,٥-٤٨,٥	٢	٥٠	١٥٠
المجموع		١٥٠		

(أ) المدى :

باعتبار أن المدى هو الفرق بين الحد الأعلى للفئة العليا ناقصاً الحد الأدنى للفئة الدنيا فهو يساوى :

٥١,٥ - ١٢,٥ = ٣٩

وباعتبار أنه الفرق بين مركزى الفئتين السالفتين :

المدى = ٥٠ - ١٤ = ٣٦

البرنامج التالى يقوم بحساب المدى لبيانات مجمعة والبيانات المستخدمة هى نفسها الواردة بالمثال (١) السابق.

```
10  REM برنامج لحساب المدي لبيانات مجمعه
20  DIM A(13),B(13),F(13),C(13)
30  T=0
40  READ N   REM عدد القيم
50  :            مركز الفئه             الفئـــــه          رقم الفئه
60  :       ____________________________________________________
70  :            ###                  ##.# - ##.#           ##
75  PRINT USING 60
80  PRINT USING 50
90  PRINT USING 60
100  PRINT
105 FOR I=1 TO N
110  READ A(I),B(I),F(I)  REM   الحد الادنى , الحد الاعلى , التكرار
120  C(I)=(A(I)+B(I))/2   REM   مركز الفئه
140  PRINT USING 70, C(I),B(I),A(I),I
150 NEXT I
160 R1=B(N)-A(1)
170 R2=C(N)-C(1)
180 PRINT ,R1;'= ١ المـدى'
190 PRINT
200 PRINT ,R2;'= ٢ المـدى'
210 PRINT
220 PRINT
230 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
240 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
250 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
260 END
```

المخرجات

رقم الفئه	الفئـــه	مركز الفئه
1	15.5 - 12.5	14
2	18.5 - 15.5	17
3	21.5 - 18.5	20
4	24.5 - 21.5	23
5	27.5 - 24.5	26
6	30.5 - 27.5	29
7	33.5 - 30.5	32
8	36.5 - 33.5	35
9	39.5 - 36.5	38
10	42.5 - 39.5	41
11	45.5 - 42.5	44
12	48.5 - 45.5	47
13	51.5 - 48.5	50

المـدى ١ = 39

المـدى ٢ = 36

(ب) الانحراف الربيعى :

كانت نتائج المثال (١٥) فى الفصل السابق كما يلى :

و٧٥٪ = ٣٤,٠٦٣

و٢٥٪ = ٢٨,٨٥٧

الانحراف الربيعى = (و٧٥٪ − و٢٥٪) / ٢ (٢)

= ٢,٦٠٣

وفيما يلى برنامج حساب الانحراف الربيعى للبيانات الواردة فى نفس المثال والمثال السابق، علماً بأن معادلة الانحراف الربيعى المستخدمة هى :

$$Y = \frac{(Q_1 - Q_3)}{2}$$

حيث :

الانحراف الربيعى = Y

الربيع الأعلى = Q_1

الربيع الأدنى = Q_3

كما أن :

$$Q(I) = L(I) + \frac{(S(I) - O(I)P(I)}{W(I)}$$

كما ورد فى برنامج حساب الربيعات فى نهاية الفصل الماضى.

```
10  REM     برنامج لحساب الانحراف الربيعى لبيانات مجمعه
20  DIM A(13),B(13),C(13),D(13),E(13),F(13),G(13)
30  T=0
40  READ N
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)    REM الحد الاعلى، الحد الادنى والتكرار
70   T=T+F(I)               REM مجموع التكرارات
80  NEXT I
90  C(1)=F(1)
100 FOR I=2 TO N
110  C(I)=C(I-1)+F(I)
120 NEXT I
130 :            التجمع التكرارى       التكرار      الفئـــه      رقم
140 :              الصاعد                                       الفئة
150 :   ------------------------------------------------------------
160 :              ###                 ##      ##.# - ##.#     ##
170 :                                  ---
180 :                                  ###                  المجموع
185 PRINT USING 150
190 PRINT USING 130
200 PRINT USING 140
210 PRINT USING 150
220 PRINT
230 FOR I=1 TO N
240  PRINT USING 160, C(I);F(I),B(I),A(I),I
250 NEXT I
260 PRINT USING 170
270 PRINT USING 180 ,T
280 PRINT
290 S1=T*3/4
300 S3=T/4
310 FOR I=1 TO N
320 IF C(I)>S1 THEN 340
330 J=I
340 NEXT I
350 FOR I=1 TO N
```

```
360 IF C(I)>S3 THEN 380
370 V=I
380 NEXT I
390 PRINT USING 430, B(J+1),A(J+1),J+1
400 PRINT
410 PRINT USING 440, B(V+1),A(V+1),V+1
420 PRINT
430 :      فئة الربيع الاعلى هى الفئة رقم ## وهى ##.# - ##.#
440 :      فئة الربيع الادنى هى الفئة رقم ## وهى ##.# - ##.#
450 GOSUB 610
460 Q1=FNA(L1,S1,O1,P1,W1)
470 PRINT ,Q1;'= الربيع الاعلى'
480 PRINT
510 GOSUB 660
520 Q3=FNA(L3,S3,O3,P3,W3)
530 PRINT ,Q3;'= الربيع الادنى'
550 PRINT
560 Y=(Q1-Q3)/2
570 PRINT ,Y;'= الانحراف الربيعى'
580 PRINT
590 PRINT
600 STOP
610 L1=A(J+1)
620 O1=C(J)
630 P1=B(J)-A(J)
640 W1=F(J+1)
650 RETURN
660 L3=A(V+1)
670 O3=C(V)
680 P3=B(V)-A(V)
690 W3=F(V+1)
700 RETURN
710 DEF FNA(L,X,O,P,W)=L+((X-O)*P)/W
720 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
730 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
740 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
750 END
```

المخرجات

رقم الفئة	الفئـــة	التكرار	التجمع التكرارى الصاعد
1	15.5 - 12.5	2	2
2	18.5 - 15.5	3	5
3	21.5 - 18.5	4	9
4	24.5 - 21.5	5	14
5	27.5 - 24.5	14	28
6	30.5 - 27.5	21	49
7	33.5 - 30.5	59	108
8	36.5 - 33.5	24	132
9	39.5 - 36.5	7	139
10	42.5 - 39.5	6	145
11	45.5 - 42.5	2	147
12	48.5 - 45.5	1	148
13	51.5 - 48.5	2	150
المجموع		150	

فئة الربيع الاعلى هى الفئة رقم 8 وهى 33.5 - 36.5

فئة الربيع الادنى هى الفئة رقم 6 وهى 27.5 - 30.5

الربيع الاعلى = 34.0625

الربيع الادنى = 28.85713

الانحراف الربيعى = 2.602684

٤ ـ الانحراف المتوسط :

هو مجموع القيم المطلقة للانحرافات عن الوسط الحسابى مقسوماً على عددها، والقيمة المطلقة (Absolute Value) هى :

$$|س| = \begin{cases} س & \text{إذا كانت } س \geqslant \text{صفراً} \\ -س & \text{إذا كانت } س < \text{صفراً} \end{cases}$$

وهـذا معناه إهمال الإشارة السالبة، واعتبار الانحراف يمثل بعداً لايهم اتجاهه. بذلك يكون :

$$\text{الانحراف المتوسط} = \frac{\sum_{ر=1}^{ن} |س_ر - \bar{س}|}{ن} \qquad (٣)$$

والسبب فى أخـذ القيمـة المـطلقة لكل انحراف هو أن مجموع الانحرافات عن الوسط الحسـابى يسـاوى صفـراً، كما سبق وأثبت فى الفصل الماضى. وبالرغم من أن الانحراف المتوسط يأخذ فى الاعتبار جميع القيم ـ وهذا ما يميزه عن المدى والانحراف الربيعى ـ فإن عمليات استخراجه شاقة، ومعادلته غير قابلة للتعامل الجبرى، إضافة إلى العيب الرئيسى وهو إهمال الإشارات، مما جعله من المقاييس غير الدقيقة. كل ذلك جعل استخدام الانحراف المتوسط فى المجالات التطبيقية يكاد يكون معدوماً.

مثال (٢ ، ٥) :

أوجد الانحراف المتوسط للقيم :

٢ ، ١ ، ٣ ، ٦ ، ٨ ، ٤

الحل:

ن = ٦ ، $\bar{س}$ = ٤

$$\text{الانحراف المتوسط} = \frac{٢ + ٣ + ١ + ٢ + ٤ + ٠}{٦}$$

= ٢

أما فى حالة التوزيعات التكرارية فالانحراف المتوسط هو :

$$\frac{|س_١ - \bar{س}|ك_١ + |س_٢ - \bar{س}|ك_٢ + |س_٣ - \bar{س}|ك_٣ + \ldots + |س_ف - \bar{س}|ك_ف}{ك_١ + ك_٢ + ك_٣ + \ldots ك_ف}$$

$$= \frac{\sum_{ر=١}^{ف} |س_ر - \bar{س}| ك_ر}{ن} \qquad (٤)$$

حيث ف هو عدد الفئات و $س_ر$ هى مراكز الفئات .

مثال (٣،٥) :

أوجد الانحراف المتوسط للبيانات التالية :

رقم الفئة	الفئة	$ك_ر$	$س_ر$	$س_ر - \bar{س}$*	$\|س_ر - \bar{س}\|ك_ر$
١	١٢,٥-١٥,٥	٢	١٤	-١٧,٤٨	٣٤,٩٦
٢	١٥,٥-١٨,٥	٣	١٧	-١٤,٤٨	٤٣,٤٤
٣	١٨,٥-٢١,٥	٤	٢٠	-١١,٤٨	٤٥,٩٢
٤	٢١,٥-٢٤,٥	٥	٢٣	-٨,٤٨	٤٢,٤٠
٥	٢٤,٥-٢٧,٥	١٤	٢٦	-٥,٤٨	٧٦,٧٢
٦	٢٧,٥-٣٠,٥	٢١	٢٩	٢,٤٨	٥٢,٠٨
٧	٣٠,٥-٣٣,٥	٥٩	٣٢	٠,٥٢	٣٠,٦٨
٨	٣٣,٥-٣٦,٥	٢٤	٣٥	٣,٥٢	٨٤,٤٨
٩	٣٦,٥-٣٩,٥	٧	٣٨	٦,٥٢	٤٥,٦٤
١٠	٣٩,٥-٤٢,٥	٦	٤١	٩,٥٢	٥٧,١٢
١١	٤٢,٥-٤٥,٥	٢	٤٤	١٢,٥٢	٢٥,٠٤
١٢	٤٥,٥-٤٨,٥	١	٤٧	١٥,٥٢	١٥,٥٢
١٣	٤٨,٥-٥١,٥	٢	٥٠	١٨,٥٢	٣٧,٠٤
	المجموع	١٥٠			٥٩١,٠٤

* $\bar{س}$ = ٣١,٤٨ من مثال (٢) فى الفصل السابق.

$$\text{الانحراف المتوسط} = \frac{\sum_{ر=١}^{١٣} \left| س_ر - \bar{س} \right| ك_ر}{ن}$$

$$\therefore \text{الانحراف المتوسط} = \frac{٥٩١,٠٤}{١٥٠}$$

$$= ٣,٩٤$$

٥ - الانحراف المعيارى :

هناك طريقة أخرى للتخلص من الإشارات السالبة للانحرافات عن الوسط الحسابى، وهى بتربيع تلك الانحرافات لتصبح جميعها موجبة ويكون مجموعها على النحو التالى :

$$\sum (س_ر - \bar{س})^٢ \qquad (٥)$$

يسمى متوسط مجموع مربعات الانحرافات المذكورة أعلاه بالتباين ($ع^٢$). إلا أنه، ولاعتبارات خاصة بالاستدلال الإحصائى، قد عدِّل تعديلاً طفيفاً ليصبح (ن - ١) بدلاً من (ن)، بذلك يكون تباين مفردات العينة هو :

$$ع^٢ = \frac{\sum_{ر=١}^{ن} (س_ر - \bar{س})^٢}{(ن - ١)} \qquad (٦)$$

افرض أن وحدة قياس القيم العينية كانت بالأمتار، إذاً فوحدة قياس الوسط الحسابى أيضاً بالأمتار. أما وحدة قياس التباين فهى الأمتار المربعة، ولكى تتوافق وحدة قياس التشتت مع وحدة قياس الوسط الحسابى والقيم العينية، فقد أخذ بالجذر التربيعى للتباين وسمى الانحراف المعيارى (ع).

وعليه يكون الانحراف المعيارى هو :

$$ع = \sqrt{\frac{\sum (س_ر - \bar{س})^٢}{ن - ١}} \qquad (٧)$$

مثال (٤ , ٥) :

أوجد التباين والانحراف المعيارى للقيم :

١٦، ٢٢، ٢١، ٢٠، ٢٣، ٢١، ١٩، ١٥، ١٣، ٢٣، ١٧، ٢٠، ٢٩، ١٨، ٢٢، ١٦، ٢٥.

الحل:

$\bar{\text{س}} = ٢٠$ (راجع مثال (١) بالفصل السابق)

$\text{ن} = ١٧$

$$\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^{٢} = (-٤)^{٤} + (٢)^{٢} + (١)^{٢} + \ldots + \ldots\ldots + (-٤)^{٢} + (٥)^{٢}$$

$$= ٢٥٤$$

$$\frac{\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^{٢}}{\text{ن} - ١} = \frac{٢٥٤}{١٦}$$

$$= ١٥,٨٧٥$$

$$\therefore \text{ع}^{٢} = ١٥,٨٧٥$$

$$\text{ع} = \sqrt{١٥,٨٧٥}$$

$$\therefore \text{ع} = ٣,٩٨٤$$

هذا وتجدر الإشارة هنا إلى أنه يمكن استخراج الانحراف المعيارى والوسط الحسابى مباشرة، باستخدام بعض الآلات الإلكترونية العلمية دون الحاجة لإجراء هذه العمليات.

كذلك يمكن إجراء بعض العمليات الجبرية على المعادلة السابقة لتصبح :

$$\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^{٢} = \sum \text{س}_{\text{ر}}^{٢} - \frac{(\sum \text{س}_{\text{ر}})^{٢}}{\text{ن}} \qquad (٨)$$

وبذلك يكون :

$$ع^{2} = \frac{\sum س_{ر}^{2} - \frac{(\sum س_{ر})^{2}}{ن}}{ن - ١} \qquad (٩)$$

فيما يلى برنـامـج لحساب التباين والانحراف المعيارى للقيم العينية الواردة فى المثال (٤) السابق علماً بأن المعادلة المستخدمة هنا هى :

$$V = \frac{T}{(N-1)}$$

حيث :

التباين = V

مجموع مربعات الانحرافات = T

عدد المتغيرات = N

وبالتالى فإن :

$$R = \sqrt{V}$$

حيث :

الانحراف المعيارى = R

```
10  REM  برنامج لحساب التباين و الانحراف المعيارى لمجموعة مفردات
20  DIM X(17)
30  S=0
40  READ N    REM  عدد القيم
50  FOR I=1 TO N
60   READ X(I)
70   S=S+X(I)
80  NEXT I
90  M=S/N
100 PRINT USING 300
105 PRINT USING 290
110 PRINT USING 300
120 PRINT
130 FOR I=1 TO N
140  D=X(I)-M
150  T=T+D**2
160  PRINT USING 310,X(I),D,D**2
170 NEXT I
180 PRINT
190 PRINT USING 320
200 PRINT USING 330,T
220 PRINT
230 V=T/(N-1)
240 R=SQR(V)
250 PRINT ,V;'=التباين'
255 PRINT
260 PRINT ,R;'=الانحراف المعيارى'
270 PRINT
280 PRINT
290 :        العدد س          س - س-          (س - س-)٢
300 :     ________________________________________________
310 :          ##               ##               ###
320 :                                          ____
330 :        المجموع                           ####
340 DATA 17,16,22,21,20,23,21,19,15,13,23,17,20,29,18,22,16,25
350 END
```

المخرجات

القيمه س	س - س-	(س - س-)٢
16	-4	16
4	2	22
1	1	21
0	0	20
9	3	23
1	1	21
1	-1	19
25	-5	15
49	-7	13
9	3	23
9	-3	17
0	0	20
81	9	29
4	-2	18
4	2	22
16	-4	16
25	5	25
المجموع		254

التباين= 15.875

الانحراف المعياري= 3.984344

أما فى حالة التوزيعات التكرارية فترجح مربعات الانحرافات بتكراراتها ليصبح على النحو التالى :

$$ع^{٢} = \frac{\sum_{ر=١}^{ف} (س_ر - \bar{س})^{٢} ك_ر}{ن - ١} \qquad (١٠)$$

وهذه أيضاً يمكن تعديلها ليكون التباين :

$$ع^{٢} = \frac{\sum س_ر^{٢} ك_ر - \frac{(\sum س_ر ك_ر)^{٢}}{ن}}{ن - ١} \qquad (١١)$$

مثال (٥,٥) :

أوجد التباين والانحراف المعيارى للبيانات الواردة أدناه :

رقم الفئة	الفئات	$ك_ر$	$س_ر$	$س_ر ك_ر$	$س_ر^{٢}$	$س_ر^{٢} ك_ر$
١	١٢,٥-١٥,٥	٢	١٤	٢٨	١٩٦	٣٩٢
٢	١٥,٥-١٨,٥	٣	١٧	٥١	٢٨٩	٨٦٧
٣	١٨,٥-٢١,٥	٤	٢٠	٨٠	٤٠٠	١٦٠٠
٤	٢١,٥-٢٤,٥	٥	٢٣	١١٥	٥٢٩	٢٦٤٥
٥	٢٤,٥-٢٧,٥	١٤	٢٦	٣٦٤	٦٧٦	٩٤٦٤
٦	٢٧,٥-٣٠,٥	٢١	٢٩	٦٠٩	٨٤١	١٧٦٦١
٧	٣٠,٥-٣٣,٥	٥٩	٣٢	١٨٨٨	١٠٢٤	٦٠٤١٦
٨	٣٣,٥-٣٦,٥	٢٤	٣٥	٨٤٠	١٢٢٥	٢٩٤٠٠
٩	٣٦,٥-٣٩,٥	٧	٣٨	٢٦٦	١٤٤٤	١٠١٠٨
١٠	٣٩,٥-٤٢,٥	٦	٤١	٢٤٦	١٦٨١	١٠٠٨٦
١١	٤٢,٥-٤٥,٥	٢	٤٤	٨٨	١٩٣٦	٣٨٧٢
١٢	٤٥,٥-٤٨,٥	١	٤٧	٤٧	٢٢٠٩	٢٢٠٩
١٣	٤٨,٥-٥١,٥	٢	٥٠	١٠٠	٢٥٠٠	٥٠٠٠
المجموع		١٥٠		٤٧٢٢		١٥٣٧٢٠

$$ع^{٢} = \frac{١٥٣٧٢٠ - \frac{٤٧٢٢ \times ٤٧٢٢}{١٥٠}}{١٤٩}$$

$$\therefore ع^{٢} = ٣٤,٠٣٦٥١$$

$$ع = ٥,٨٣٤$$

يتضح من المعادلات السابقة أن قيمة التباين لا تكون إلا موجبة ؛ لأن البسط عبارة عن مجموع مربعات، والمربعات لاتكون إلا موجبة ؛ لذا فإن أقل قيمة للتباين هى الصفر. وهذه لا تتحقق إلا فى حالة واحدة، وهى عندما تكون القيم متساوية تماماً، وهذا يعنى التجانس التام بين القيم، أما أعلى قيمة له فلا حدود لها.

هذا وتجدر الإشارة هنا إلى أن التباين يتناقص مع زيادة حجم العينة (ن)، كما هو واضح من المعادلات السالفة الذكر.

لحساب التباين والانحراف المعيارى لبيانات مبوبة، فإننا نستخدم البرنامج التالى للبيانات الواردة بالمثال (٥) وباستخدام المعادلة :

$$V = \frac{\left(T_3 - \frac{(T_2)^2}{T_1} \right)}{T_1 - 1}$$

حيث :

V = التباين

T_3 = مجموع مربعات مراكز الفئات مرجحة بتكراراتها

T_2 = مجموع مراكز الفئات مرجحة بتكراراتها

T_1 = مجموع التكرارات

وكذلك

$$R = \sqrt{V}$$

حيث :

R = الانحراف المعيارى

```
10  REM   برنامج لحساب التباين والانحراف المعيارى لبيانات مبوبه
20  DIM A(13),B(13),C(13),D(13),E(13),F(13),G(13)
30  F1=0
40  READ N     REM  عدد المشاهدات
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)   REM  الحد الادنى , الحد الاعلى , التكرار
70  NEXT I
80  PRINT USING 320
85  PRINT USING 300
90  PRINT USING 310
100 PRINT USING 320
110 PRINT
120 FOR I= 1 TO N
130  C(I)=(A(I)+B(I))/2
140  G(I)=C(I)*F(I)
150  D(I)=C(I)**2
160  E(I)=D(I)*F(I)
170  T1=T1+F(I)
180  T2=T2+G(I)
190  T3=T3+E(I)
200  PRINT USING 330, E(I),D(I),G(I),C(I),F(I),B(I),A(I),I
210 NEXT I
220 PRINT USING 340
230 PRINT USING 350,T3,T2,T1
240 V=(T3-T2**2/T1)/(T1-1)
250 R=SQR(V)
260 PRINT
270 PRINT
280 PRINT ,V;'=التباين'
285 PRINT
290 PRINT ,R;'=الانحراف المعيارى'
300 :   س ر٢ ك    س ر٢    س ر ك    س ر    ك    الفئـــه    رقم الفئه
310 :
320 : ----------------------------------------------------------------
330 :   #####    ####    ####    ##    ##    ##.# - ##.#    ##
340 :  ______          ____          ___
350 :  ######          ####          ###                 المجموع
360 PRINT
370 PRINT
380 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
390 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
400 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
410 END
```

المخرجات

س ر٢ ك	س ر٢	س ر ك	س ر	ك	الفئـــه	رقم الفئه
392	196	28	14	2	15.5 - 12.5	1
867	289	51	17	3	18.5 - 15.5	2
1600	400	80	20	4	21.5 - 18.5	3
2645	529	115	23	5	24.5 - 21.5	4
9464	676	364	26	14	27.5 - 24.5	5
17661	841	609	29	21	30.5 - 27.5	6
60416	1024	1888	32	59	33.5 - 30.5	7
29400	1225	840	35	24	36.5 - 33.5	8
10108	1444	266	38	7	39.5 - 36.5	9
10086	1681	246	41	6	42.5 - 39.5	10
3872	1936	88	44	2	45.5 - 42.5	11
2209	2209	47	47	1	48.5 - 45.5	12
5000	2500	100	50	2	51.5 - 48.5	13
153720		4722		150		المجموع

34.03691 =التباين

5.834116 =الانحراف المعيارى

٦ـ الانحراف المعياري والمقارنات :

يحتاج المرء كثيراً لإجراء المقارنة بين تشتتي مجموعتين مختلفتين في وسطيهما وانحرافيهما. وربما تكونان مختلفتين حتى في وحدتي القياس، كالأعمار والأجور ـ مثلاً ـ وربما يبدو لأول وهلة أن المجموعة ذات الانحراف المعياري الأكبر هي الأكثر تشتتاً. ولكن هب أن الوسط الحسابي لمجموعة ما كان بالكيلومترات، وكذلك الانحراف المعياري، فإذا تم تحويل الوسط والانحراف إلى أمتار، ضرب كل منهما في ألف، فأصبح الانحراف المعياري يساوي ألف مرة، على ما كان عليه في الحالة الأولى، فهل هذا يعني أن تشتت هذه المجموعة قد ازداد؟

إذاً فالانحراف المعياري، وكذلك الوسط الحسابي، يتأثران بوحدات القياس؛ لذلك لابد من اللجوء إلى مقياس آخر يخلو من وحدات القياس. وهذا ما توصل إليه كارل بيروسون (١٨٥٧-١٩٣٦) عندما أثبت أن نسبة الانحراف المعياري إلى الوسط الحسابي لنفس المجموعة تمثل مقياساً أفضل لمقارنة التشتت بين المجموعتين، ولقد سمى هذا المقياس بمعامل الاختلاف (Coefficient of Variation) .

إذاً

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times ١٠٠$$

أي أن :

$$\text{معامل الاختلاف} = \frac{ع}{\bar{س}} \times ١٠٠ \qquad (١٢)$$

ومعامل الاختلاف نسبة مئوية تخلو خلواً تاماً من وحدات القياس، ويمكن استخدامه لمقارنة التشتت بين أي مجموعتين، سواء بنفس وحدة القياس أم بغيرها.

مثال (٦ , ٥) :

الوسط الحسابي لمجموعة ما يساوي ١٥٠٠، بينما كان الانحراف المعياري لنفس المجموعة وبنفس وحدة القياس يساوي ١٠٠ . أما الوسط الحسابي لمجموعة ثانية وبوحدة قياس مختلفة فيساوي ٦٠، وانحرافها المعياري يساوي ٥ . فأي المجموعتين أكثر تشتتاً؟

الحل :

الانحراف المعيارى للمجموعة الأولى يساوى عشرين مثلاً للانحراف المعيارى الخاص بالمجموعة الثانية ، إلا أن ذلك ليس دليلاً على أن المجموعة الأولى هى الأكثر تشتتاً، إذ لابد من استخدام معامل الاختلاف.

معامل الاختلاف للمجموعة الأولى = $\frac{١٠٠}{١٥٠}$ × ١٠٠

= ٦,٦٦٧٪

معامل الاختلاف للمجموعة الثانية = $\frac{٥}{٦٠}$ × ١٠٠

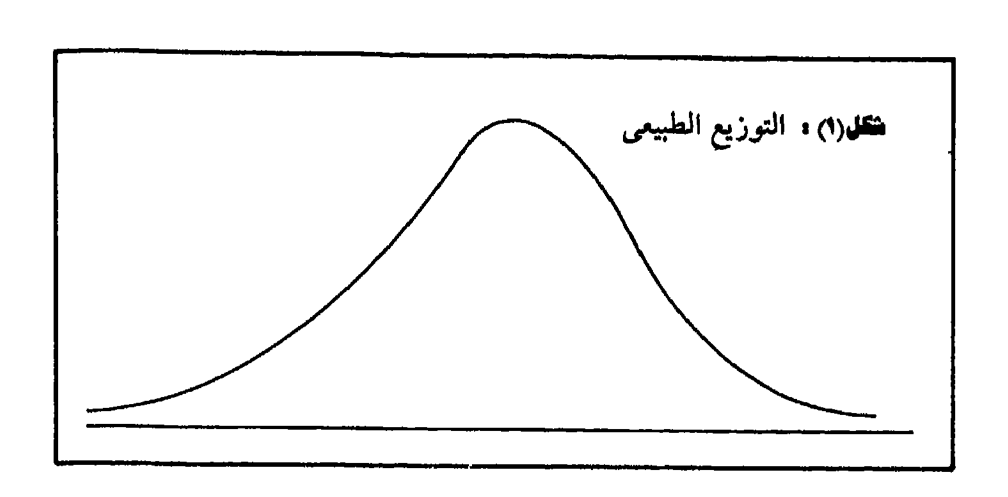
= ٨,٣٣٪

إذاً فالمجموعة الثانية هى الأكثر تشتتاً.

أما إذا كان الهدف هو إجراء المقارنة بين قيمتين تنتمى كل منهما إلى مجموعة مختلفة ، كمقارنة درجات الطالب فى مواد مختلفة ، أو مقارنة درجات طالبين فى مجموعتين مختلفتين لنفس المواد ، سواء كانت القيم بنفس وحدة القياس أو بوحدتى قياس مختلفتين، فلا بد من استخدام وحدات قياس متناظرة بين المجموعتين.

بيد أن وحدة القياس الجديدة يجب أن تخلو أيضاً من وحدات القياس ، وعليه فلا بد أن تمثل النسبة بين القيم وانحرافاتها المعيارية ، إلا أن طبيعة القيم قد تكون أصلاً كبيرة فى مجموعة ، وصغيرة فى مجموعة أخرى؛ لذلك لا بد من أخذ الوسطين الحسابيين فى الاعتبار. ويسمى المقياس الناتج بعد إجراء هذه المقارنات بالقيم المعيارية (Standardized value) . وتحتاج القيم المعيارية لتوفير شروط معينة قبل تطبيقها وتعريفها ، ويمكن إيجاز هذه الشروط فى الفقرات التالية :

يقترب التوزيع الخاص بالبيانات من التوزيع الطبيعى كلما ازداد حجم العينة (ن) إلى أن يصبح توزيعاً طبيعياً ، عندما يكون حجم العينة كبيراً . والتوزيع الطبيعى هو توزيع متماثل ، يتساوى فيه الوسط والوسيط والمنوال ، وهو أشبه بالناقوس المقلوب مثلما هو مبين فى الشكل التالى.

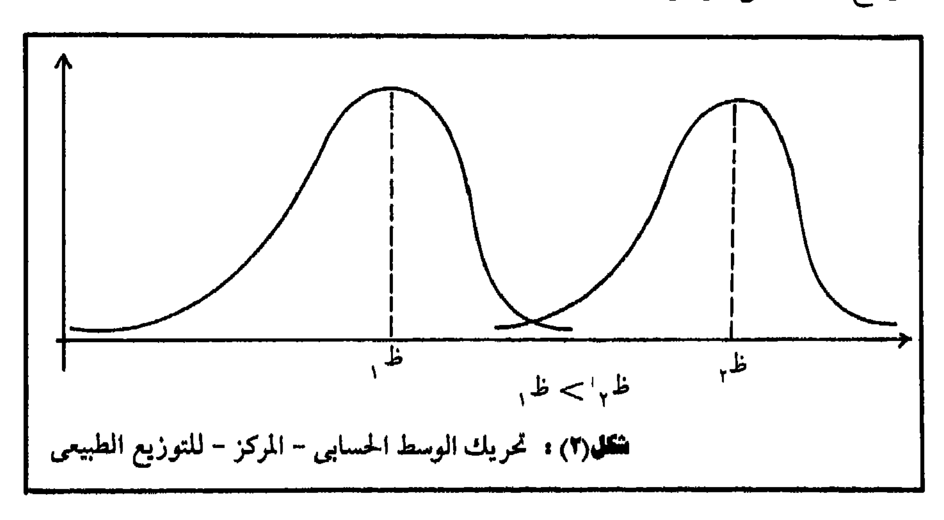


شكل(١) : التوزيع الطبيعى

وبما أن منحنى التوزيع الطبيعى هو أحد المنحنيات الاحتمالية المتصلة، بل هو أشهرها، فإن المساحة الواقعة تحته تساوى واحداً. هذا ويمتد منحنى التوزيع الطبيعى من

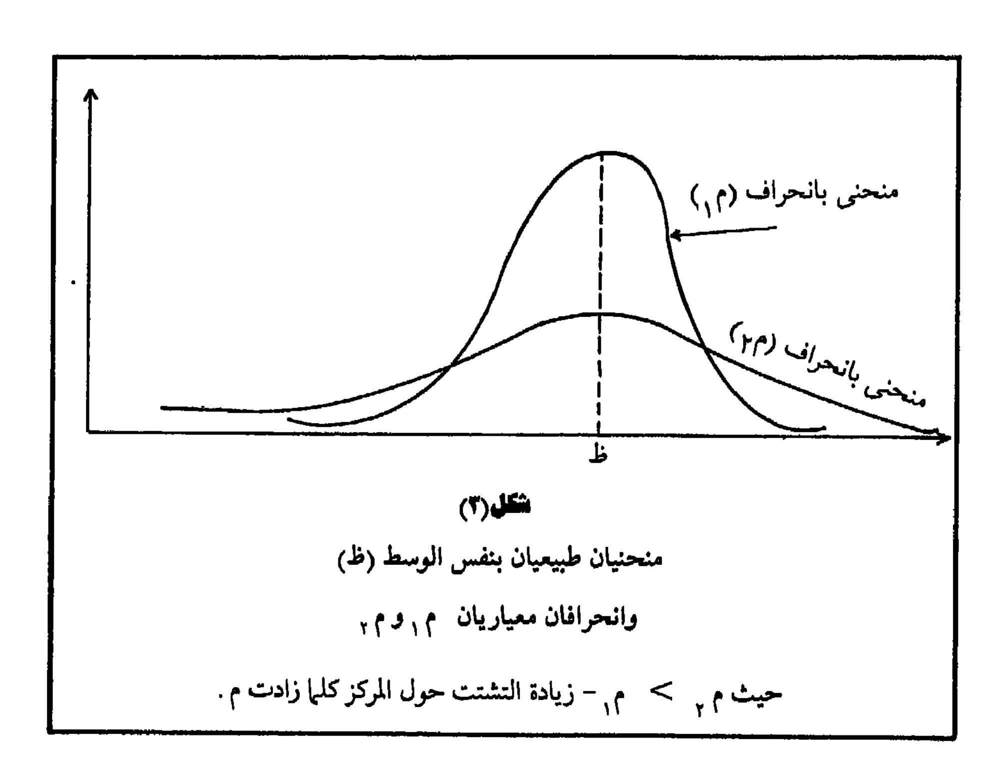
- ∞ إلى + ∞ ويتميز بطرفين رفيعين للغاية يوشكان ملامسة المحور الأفقى.

إذا كانت ظ هى الوسط الحسابى للمجتمع الذى سحبت منه العينة، وإذا كانت م هى الانحراف المعيارى لبيانات المجتمع، فهذا يعنى أن سَ و ع هما مقدران لقيمتى ظ و م. أما ظ و م فتسميان مَعْلَمَى التوزيع الطبيعى إذا كان التوزيع طبيعياً حقاً، وعندها يكون التوزيع متماثلاً حول الخط العمودى الساقط على ظ، فإذا تحركت ظ إلى اليمين أو اليسار انتقل معها التوزيع؛ لأنها تمثل المرتكز بالنسبة له.



شكل(٢) : تحريك الوسط الحسابى - المركز - للتوزيع الطبيعى

أما المعلم الثاني (م) فهو مقياس التشتت؛ لذلك فهى التى تحدد شكل المنحنى بعد أن يكون موقعه قد حدِّد بواسطة (ظ). إذ كلما قلت م زاد ارتفاع المنحنى وقل تشتته، والعكس صحيح.



شكل(٣)

منحنيان طبيعيان بنفس الوسط (ظ)

وانحرافان معياريان م١ و م٢

حيث م٢ > م١ - زيادة التشتت حول المركز كلما زادت م.

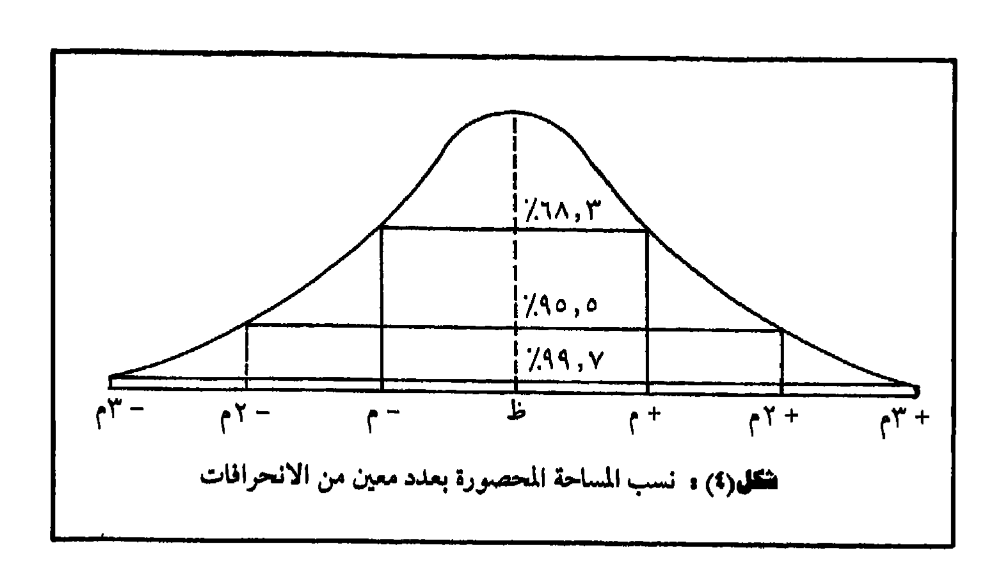
يمكن تقسيم المساحة الواقعة تحت المنحنى الطبيعى بعدد الانحرافات الواقعة إلى يمين أو يسار الوسط، إذ أن :

(أ) ٦٨,٣٪ من المساحة تنحصر فى الفترة ظ + م و ظ − م.

(ب) ٨٦,٦٪ من المساحة تنحصر فى الفترة ظ + ١ $\frac{1}{2}$ م و ظ − ١ $\frac{1}{2}$ م.

(ج) ٩٥,٥٪ من المساحة تنحصر فى الفترة ظ + ٢ م و ظ − ٢ م.

(د) ٩٩,٧٪ من المساحة تنحصر فى الفترة ظ + ٣ م و ظ − ٣ م.



شكل(٤) : نسب المساحة المحصورة بعدد معين من الانحرافات

أما التوزيع الناتج من توزيعين طبيعيين أو أكثر، فهو توزيع طبيعى أيضاً، فالتوزيع المشترك لمجموع مجتمعين مستقلين عدد عناصر كل منهما يساوى (ن)، هو توزيع طبيعى بوسط حسابى يساوى مجموع الوسطين وتباين يساوى مجموع التباينين أيضاً. أما التوزيع المشترك للفرق بين التوزيعين فهو أيضاً طبيعى بوسط حسابى يساوى الفرق بين الوسطين وبتباين يساوى مجموعيهما. هذا وتجدر الإشارة هنا إلى أن التباين يجمع ولا يطرح؛ لأن تشتت التوزيع المشترك يجمع بين التشتتين فى الحالتين معاً.

وأما إذا كانت س هى متغير عشوائى تتبع توزيعاً طبيعياً، وإذا كانت ى هى متغير جديد يمكن التعبير عنه بالمعادلة :

$$ى = \frac{س - ظ}{م} \qquad (١٣)$$

فالتوزيع الخاص بالمتغير ى هو توزيع طبيعى أيضاً بوسط حسابى يساوى صفراً، وانحراف معيارى واحد. ويسمى هذا التوزيع بالتوزيع الطبيعى المعيارى (Standard Normal Distrib.)

إذاً فلكل قيمة من س ما يناظرها فى ى. فالقيمة س ر التى تنتمى لعينة ذات توزيع طبيعى بوسط حسابى (س̄) وانحراف معيارى (ع) يمكن التعبير عنها بالقيمة المعيارية :

$$ى = \frac{س_ر - \bar{س}}{ع} \quad (١٤)$$

والقيمة المعيارية هى التى تستخدم للمقارنة بين القيم التى تتبع لمجتمعات مختلفة على أساس عدد الوحدات المعيارية الناتجة بعد التحويل، والتى توضح الترتيب الخاص بكل متغير فى مجتمعه اعتماداً على الشكل (٤).

مثال (٥,٧) :

حصل أحد الطلاب على ٨١ درجة فى أحد الامتحانات التى كان الوسط العام فيها لجميع الطلاب المتحنين ٧٠ درجة بانحراف معيارى قدره ١٠، بينما حصل طالب آخر فى مؤسسة تعليمية أخرى على ٩٠ درجة فى نفس المادة، وكان الوسط الحسابى والانحراف المعيارى للممتحنين فى المدرسة الثانية هما ٧٥ و ١٦ على التوالى. فأى الطالبين أفضل، مقارنة بالمجموعة التى يدرس معها؟

الحل :

$$ى_١ = \frac{٨١ - ٧٠}{١٠}$$

$$= ١,١ \quad \text{انحراف معيارى}$$

$$ى_٢ = \frac{٩٠ - ٧٥}{١٦}$$

$$= \frac{١٥}{١٦}$$

$$= ٠,٩٤ \quad \text{انحراف معيارى}$$

إذاً فمستوى الطالب الأول هو الأفضل؛ لأنه يبعد عن الوسط بـ ١,١ انحراف معيارى، بينما يزيد الثانى على الوسط بمقدار ٠,٩٤ انحراف معيارى.

البرنامج التالى يقوم بحساب القيمة المعيارية لأى رقم من مجموعة أرقام باستخدام المعادلة :

$$Y = \frac{X - M}{R}$$

حيث :

Y	=	القيمة المعيارية
X	=	المتغير
M	=	الوسط الحسابى
R	=	الانحراف المعيارى

```
10  REM برنامج لحساب القيمه المعياريه لمجموعه مفردات
20  DIM X(17)
30  S=0
40  READ N   REM عدد القيم
50  PRINT USING 330
60  PRINT USING 340
70  FOR I=1 TO N
80   READ X(I)
90   PRINT USING 350,X(I)
100  S=S+X(I)
110 NEXT I
120 PRINT USING 360
130 PRINT USING 370,S
140 M=S/N
150 PRINT
160 FOR I=1 TO N
170  D=X(I)-M
180  T=T+D**2
190 NEXT I
200 PRINT
210 PRINT
220 V=T/(N-1)
230 R=SQR(V)
240 PRINT ,M;'=الوسط الحسابى'
250 PRINT
260 PRINT ,R;'=الانحراف المعيارى'
270 PRINT
280 Y=(X(5)-M)/R
290 PRINT USING 380,Y,X(5)
300 PRINT
310 PRINT
330 :                              القيم
340 :                              _____
350 :                                ##
360 :                               ____
370 :                               ###                 المجموع
380 :                    ##.### هى ## القيمه المعياريه للرقم
390 DATA 17,16,22,21,20,23,21,19,15,13,23,17,20,29,18,22,16,25
400 END
```

المخرجات

القيم	
16	
22	
21	
20	
23	
21	
19	
15	
13	
23	
17	
20	
29	
18	
22	
16	
25	
340	المجموع

الوسط الحسابى= 20

الانحراف المعيارى= 3.984344

القيمه المعياريه للرقم 23 هى 0.753

٧ ـ العزوم (Moments)

إذا كانت $س_١$ ، $س_٢$ ، $س_٣$ ، . . . . . ، $س_ن$ ، قيماً عينية ، فالعزم (ل) للدرجة (و) حول النقطة (أ) هو الإحصائية :

$$ل_و = \frac{١}{ن} \sum_{ر=١}^{ن} (س_ر - أ)^و \qquad (١٥)$$

فالعزم من الدرجة (و) حول نقطة الأصل (أ = صفراً) هو :

$$ل_و = \frac{١}{ن} \sum س_ر^و \qquad (١٦)$$

أما العزم من الدرجة (و) حول الوسط الحسابى (أ = $\bar{س}$) فهو :

$$ل_و = \frac{١}{ن} \sum (س_ر - \bar{س})^و \qquad (١٧)$$

أمـا فى حالـة التـوزيعـات التكرارية فترفع الانحرافات إلى درجة العزم، ثم تضرب فى التكرار، كما هو الحال فى الوسط الحسابى والتباين. إذاً فالعزم حول نقطة الأصل لبيانات مبوبة هو :

$$\text{ل}_{\text{و}} = \frac{1}{\text{ن}} \sum_{\text{ر}=1}^{\text{ف}} \text{س}_{\text{ر}}^{\text{و}}\ \text{ك}_{\text{ر}} \qquad (١٨)$$

والعزم حول الوسط الحسابى للبيانات المبوبة أيضاً هو :

$$\text{ل}_{\text{و}} = \frac{1}{\text{ن}} \sum_{\text{ر}=1}^{\text{ف}} (\text{س}_{\text{ر}} - \bar{\text{س}})^{\text{و}}\ \text{ك}_{\text{ر}} \qquad (١٩)$$

يكون العزم من الدرجة صفر إذا كانت و = صفراً ومن الدرجة الأولى إذا كانت و = ١ ومن الثانية إذا كانت و = ٢ وهكذا. كذلك يؤخذ العزم فى أكثر الحالات حول الوسط الحسابى. لذلك يجب أن تحدد أى نقطة أخرى اعتبرت مرتكزاً للعزم بخلاف الوسط الحسابى.

يلاحظ أن العزم من الدرجة صفر حول أى نقطة يساوى واحداً؛ ذلك لأن :

$$\text{ل}_{0} = \frac{1}{\text{ن}} \sum (\text{س}_{\text{ر}} - \text{أ})^{0} \qquad (٢٠)$$

$$\therefore \text{ل}_{0} = \frac{1}{\text{ن}} \sum 1$$

$$= \frac{\text{ن}}{\text{ن}}$$

$$\therefore \text{ل}_{0} = 1$$

أما العزم الأول حول نقطة الأصل فيساوى الوسط، بينما يمثل العزم الثانى حول نقطة الأصل أيضاً متوسط مجموع مربعات المتغيرات، ذلك لأن :

$$\text{ل}_{1} = \frac{1}{\text{ن}} \sum \text{س}_{\text{ر}}$$

$$= \bar{\text{س}}$$

$$ل_٢ = \frac{١}{ن} \sum س_ر^٢$$

وبذلك يكون التباين هو :

$$ع^٢ = \frac{ن}{ن-١} (ل_٢ - ل_١^٢) \qquad (٢١)$$

أما العزم الأول (حول الوسط الحسابى) فيساوى صفراً، والعزم الثانى هو متوسط مربعات الانحرافات عن الوسط الحسابى . أى أن :

$$ل_١ = \frac{١}{ن} \sum (س_ر - \bar{س}) \qquad (٢٢)$$

$$= صفراً$$

$$ل_٢ = \frac{١}{ن} \sum (س_ر - \bar{س})^٢ \qquad (٢٣)$$

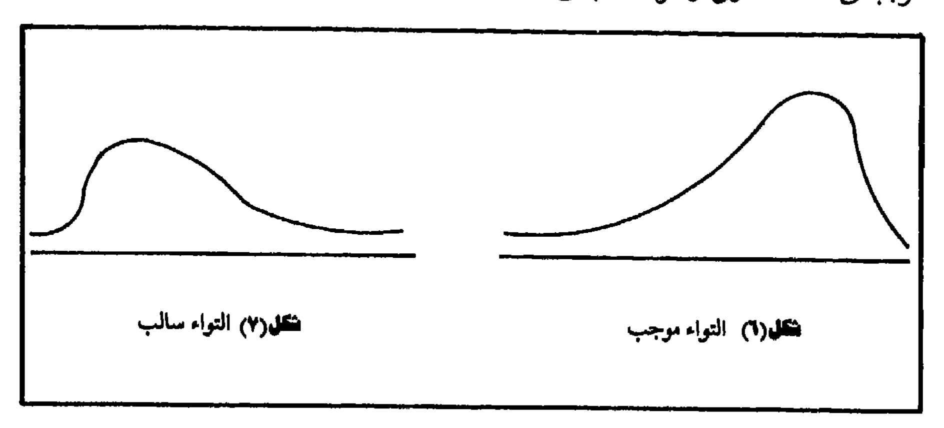
لذلك يعتبر العزم الثانى هو التباين؛ وذلك لضآلة الفرق بين قيمة ن وقيمة (ن - ١)، خاصة عندما يكون حجم العينة كبيراً.

وتجدرالإشارة هنا إلى أن كلمة العزم قد وفدت إلى الإحصاء من مجال الفيزياء، فالعزم فى الفيزياء هو ناتج ضرب القوة العمودية فى طول الذراع ، فالقوة العمودية هنا هى التكرار، أما طول الذراع فهو المسافة بين نقطة المرتكز ($\bar{س}$) والنقطة المعينة ($س_ر$).

اتضح مما مضى أن العزمين الأول والثانى حول نقطتى الأصل والوسط الحسابى يستخدمان لاستخراج الوسط الحسابى والتباين، وللتحقق من أن مجموع الانحرافات حول الوسط الحسابى يساوى صفراً. أما العزم الثالث والعزم الرابع حول الوسط الحسابى فيستخدمان لاستخراج معاملى الالتواء والتفرطح التالى ذكرهما على التوالى. هذا وتجدر الإشارة هنا إلى أن العزوم الفردية - $ل_١$، $ل_٣$، $ل_٥$، ... - تساوى صفراً فى حالة التوزيع الطبيعى المتماثل، وتقترب من الصفر كلما اقترب توزيع البيانات من التماثل، أما العزوم الزوجية فهى دائماً موجبة ولا تساوى صفراً، إلا إذا كانت جميع القيم متساوية.

٨ - الالتواء (Skewness)

ورد فى مقارنة الوسط الحسابى، والوسيط، والمنوال أن الوسيط يكون دائماً بين الوسط والمنوال. فإذا كانت هناك قيم كبيرة لدرجة التطرف فالوسط الحسابى هو الأكبر؛ لأن التوزيع يكون ملتوياً فى اتجاه البيانات الكبرى - ناحية اليمين. أما إذا كانت هناك قيم صغيرة جداً، فيكون الوسط الحسابى هو الأصغر، ويكون الالتواء للجهة اليسرى. فيقال إن هناك التواءً موجباً فى الحالة الأولى والتواء سالباً فى الحالة الثانية.


شكل(٦) التواء موجب

شكل(٧) التواء سالب

إذاً فالالتواء هو عدم تماثل التوزيع بالنسبة لأى خط عمودى، لذلك يتم قياس الالتواء لتحديد ما إذا كان هناك التواء موجب أو سالب ودرجة الالتواء لتوضيح شكل منحنى توزيع البيانات، ويستخدم فى ذلك العزم الثالث.

فإذا تساوى مجموع مكعبات الانحرافات الخاصة بالقيم التى تقل عن الوسط الحسابى بمجموع مكعبات انحرافات القيم التى تزيد على الوسط الحسابى ، يكون العزم الثالث صفراً. إذاً فمعامل الالتواء للتوزيع المتماثل يساوى صفراً، طالما أنه يعتمد على العزم الثالث :

$$ل_٣ = \frac{\sum (س_ر - \bar{س})^٣}{ن} \qquad (٢٤)$$

يكون العزم الثالث موجباً إذا كان الالتواء للجهة اليمنى ، وسالباً إذا كان الالتواء للجهة اليسرى، ويبتعد عن الصفر كلما كان الالتواء شديداً. ذلك لأن مجموع المكعبات الموجبة يكون أكثر من مجموع المكعبات السالبة فى حالة الالتواء الموجب، والعكس صحيح فى حالة الالتواء السالب.

بيد أن العزم الثالث يتأثر كثيراً بوحدة القياس أو مقياس الرسم ؛ لذلك لا يمكن استخدامه للمقارنة بين التواءين. ولقد استخدمت القيمة المعيارية للالتواء وسميت بمعامل الالتواء (Coefficient of Skewness) للتخلص من وحدات القياس ؛ لهذا فقد عرف معامل الالتواء بأنه :

$$\sqrt{ت_1} = \frac{ل_3}{ع^3} \qquad (٢٥)$$

ويلاحظ أن المقام هو مكعب الانحراف المعيارى حتى يخلو معامل الالتواء خلواً تاماً من وحدات القياس أو مقياس الرسم. كذلك يمكن تعريف معامل الالتواء بأنه :

$$ت_1 = \frac{(ل_3)^2}{(ع^3)^2} \qquad (٢٦)$$

وبما أن $ع = \sqrt{ل_2}$ فإن :

$$(ع^3)^2 = ل_2^3 \qquad (٢٧)$$

وبذلك تكون :

$$ت_1 = \frac{ل_3^2}{ل_2^3} \qquad (٢٨)$$

$$= \frac{\text{مربع العزم الثالث}}{\text{مكعب العزم الثانى}}$$

كما أن :

$$\sqrt{ت_1} = \frac{ل_3}{ل_2 \sqrt{ل_2}} \qquad (٢٩)$$

مثال (٨,٥) :

أوجد معامل الالتواء للتوزيع التالى :

الفئات	ك ر	س ر	س ر ك ر	س ر - سَ	(س ر - سَ)^{2}	(س ر - سَ)^{2} ك ر	(س ر - سَ)^{3} ك ر
٧-٣	١	٥	٥	-١٢,٢٤	١٤٩,٨١٧٦	١٤٩,٨١٧٦	-١٨٣٣,٧٦٧٤
١١-٧	٣	٩	٢٧	-٨,٢٤	٦٧,٨٩٧٦	٢٠٣,٦٩٢٨	-١٦٧٨,٤٢٨٦
١٥-١١	١١	١٣	١٤٣	-٤,٢٤	١٧,٩٧٧٦	١٩٧,٧٥٣٦	-٨٣٨,٤٧٥٢٦
١٩-١٥	٢٠	١٧	٣٤٠	-٠,٤	٠,٠٥٧٦	١,١٥٢	-٠,٢٧٦٤٨
٢٣-١٩	٩	٢١	١٨٩	٣,٧٦	١٤,١٣٧٦	١٢٧,٢٣٨٤	٤٧٨,٤١٦٣٨
٢٧-٢٣	٤	٢٥	١٠٠	٧,٧٦	٦٠,٢١٧٦	٢٤٠,٨٧٠٤	١٨٦٩,١٥٤٢
٣١-٢٧	٢	٢٩	٥٨	١١,٧٦	١٣٨,٢٩٧٦	٢٧٦,٥٩٥٢	٣٢٥٢,٧٥٩٦
المجموع	٥٠		٨٦٢		٤٤٨,٤٠٣٢	١١٩٧,١٢٠	١٢٤٩,٣٨٢٤

سَ = ١٧,٢٤

$$\text{ل}_{2} = \frac{\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^{2}\ \text{ك}_{\text{ر}}}{\text{ن}}$$

$$= \frac{١١٩٧,١٢٠}{٥٠}$$

$$= ٢٣,٩٤٢$$

$$\text{ل}_{3} = \frac{\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^{3}\ \text{ك}_{\text{ر}}}{\text{ن}}$$

$$= \frac{١٢٤٩,٣٨٢٤}{٥٠}$$

$$\therefore \quad ل_٣ = ٢٤,٩٨٨$$

$$\sqrt{ت_١} = \frac{ل_٣}{ل_٢ \sqrt{ل_٢}}$$

$$= \frac{٢٤,٩٨٨}{٢٣,٩٤٢ \sqrt{٢٣,٩٤٢}}$$

$$\therefore \sqrt{ت_١} = ٠,٢١٣$$

$$ت_١ = ٠,٠٤٥٥$$

البرنامج التالى يحسب معامل الالتواء والذى سبق الحديث عنه مستخدماً فى ذلك التوزيع التكرارى الوارد بالمثال (٨) السابق باستخدام المعادلة :

$$W = Q^2$$

حيث :

$$Q = \frac{L_3}{L_2 \sqrt{L_2}}$$

L_3 = العزم الثالث

L_2 = العزم الثانى

```
10  REM   برنامج لحساب الالتـــواء
20  DIM A(7),B(7),C(7),D(7),E(7),F(7),G(7),H(7),K(7),L(7)
30  P=0
40  V=0
50  R=0
60- S=0
70  T=0
75  D1=0
80  READ N     REM  عدد المشاهدات
90  FOR I=1 TO N
100  READ A(I),B(I),F(I)   REM  الحد الادنى , الحد الاعلى , التكرار
110  T=T+F(I)
120  C(I)=(A(I)+B(I))/2
130  D(I)=C(I)*F(I)
140  D1=D1+D(I)
150 NEXT I
160 M=D1/T
170 FOR I=1 TO N
180 E(I)=C(I)-M
190 K(I)=(E(I))**2
200 L(I)=(C(I)-M)**2*F(I)
210 G(I)=(C(I)-M)**3*F(I)
220 H(I)=C(I)*F(I)
230 P=P+F(I)
240 V=V+H(I)
250 R=R+K(I)
260 S=S+L(I)
270 U=U+G(I)
280 NEXT I
285 PRINT USING 305
290 PRINT USING 300
295 PRINT USING 305
300 : الفئـــــة  كر  ســر  سركر  (سر-س_)  (سر-س_)٢  (سر-س_)٢كر  (سر-س_)٣كر
305 : ____________________________________________________________________
310 PRINT
320 FOR I=1 TO N
330 PRINT USING 340, G(I),L(I),K(I),E(I),H(I),C(I),F(I),B(I),A(I)
340 :   ######.###    ####.###    ###.###   ###.###  ###   ###  ##  ##.# - ##.=
350 NEXT I
360 PRINT USING 370
370 : ___________    _________    ________          ______     ___
380 PRINT
390 PRINT USING 400, U,S,R,V,P
400 :  #####.####   #####.###    ####.###          ####       ###  المجموع
401 L2=S/P
402 L3=U/P
403 Q=L3/(L2*SQR(L2))
405 W=Q**2
410 PRINT
420 PRINT USING 430, W
430 :                    ##.#### =معامل الالتواء
440 PRINT
450 PRINT
460 DATA 7,3,7,1,7,11,3,11,15,11,15,19,20,19,23,9,23,27,4,27,31,2
470 END
```

المخرجات

(س ر -س_)٣كر	(س ر -س_)٢كر	(س ر -س_)٢	(س ر -س_)	س ر كر	س ر	كر	الفئة
-1833.763	149.817	149.817	-12.240	5	5	1	7.0 - 3.0
-1678.422	203.692	67.897	-8.240	27	9	3	11.0 - 7.0
-838.469	197.753	17.978	-4.240	143	13	11	15.0 - 11.0
-0.276	1.152	0.058	-0.240	340	17	20	19.0 - 15.0
478.420	127.239	14.138	3.760	189	21	9	23.0 - 19.0
1869.160	240.871	60.218	7.760	100	25	4	27.0 - 23.0
3252.767	276.595	138.298	11.760	58	29	2	31.0 - 27.0
1249.4210	1197.119	448.403		862		50	المجموع

0.0455 =معامل الالتواء

لتحديد ما إذا كان التوزيع ملتوياً بدرجة كبيرة (جوهرية) أم لا، يمكن مقارنة قيمة √تَ بالجدول (م١١) الملحق بنهاية الفصل. فإذا كانت القيمة الموجبة تزيد على القيمة المناظرة لها بالجدول تحت العمود الخاص بنسبة ٥٪، أو كانت القيمة المحسوبة سالبة وأقل من القيمة المبينة فى الملحق، فالالتواء بدرجة كبيرة. ويلاحظ هنا أن قيمة √تَ هنا موجبة، ولكنها أقل من القيمة المناظرة لها فى الجدول، والتى تساوى ٠,٥٣٤ وعليه يكون الالتواء ظاهرياً، أى أن توزيع البيانات لايختلف كثيراً عن التوزيع الطبيعى.

أما فى حالة التوزيعات ذات الفئات المفتوحة فيمكن استخدام الربيعات والمدى الربيعى لتقدير الالتواء. فالفرق بين الوسيط والربيع الأعلى يساوى الفرق بين الوسيط والربيع الأدنى فى حالة التماثل، أى أن :

و٧٥٪ - و٢٥٪ = و٥٠٪ - و٢٥٪ فى حالة التماثل

وأما إذا كان الالتواء سالباً فيكون الوسيط أقرب للربيع الأعلى، والعكس صحيح. لذلك يعرف معامل الالتواء فى هذه الحالة بأنه :

تَ١ = [(و٧٥٪ - و٥٠٪) - (و٥٠٪ - و٢٥٪)] ÷ (و٧٥٪ - و٢٥٪) (٣٠)

= (و٧٥٪ - ٢و٥٠٪ + و٢٥٪) ÷ (و٧٥٪ - و٢٥٪)

يستخدم البعض نسبة الفرق بين الوسط الحسابى والوسيط أو المنوال إلى الانحراف المعيارى، كمقياس لمعامل الالتواء، إلا أن هذه الطريقة ليست دقيقة؛ لأن الانحراف المعيارى مرتبط بالوسط الحسابى أكثر من ارتباطه ببقية مقاييس النزعة المركزية، كما أن المنوال مقياس غير دقيق فى حد ذاته.

مثال (٥،٩) :

استخدم البيانات التالية لتقدير الالتواء بطريقة الربيعات :

الفئات	ك ر	تجمع تكرارى صاعد
٣ـ٧	١	١
٧ـ١١	٣	٤
١١ـ١٥	١١	١٥
١٥ـ١٩	٢٠	٣٥
١٩ـ٢٣	٩	٤٤
٢٣ـ٢٧	٤	٤٨
٢٧ـ٣١	٢	٥٠
المجموع	٥٠	

$$\text{و}_{\text{ن}}\% = \text{ح}_{1} + \frac{\left(\frac{\text{ر ن}}{\text{١٠٠}} - \text{ن}'\right)\text{ط}}{\text{ك}}$$

$$\text{و}_{\text{٥٠}}\% = \text{١٥} + \frac{(\text{٢٥} - \text{١٥}) \times \text{٤}}{\text{٢٠}}$$

$$= \text{١٧}$$

$$\text{و}_{\text{٢٥}}\% = \text{١١} + \frac{(\text{١٢,٥} - \text{٤}) \times \text{٤}}{\text{١١}}$$

$$= \text{١٤,٠٩}$$

$$\text{و}_{\text{٧٥}}\% = \text{١٩} + \frac{(\text{٣٧,٥} - \text{٣٥}) \times \text{٤}}{\text{٩}}$$

$$= \text{٢٠,١١}$$

$$ت_١ = \frac{٢٠,١١-٢ \times ١٧ + ١٤,٠٩}{٢٠,١١ - ١٤,٠٩}$$

$$\therefore ت_١ = ٠,٠٣٣$$

البرنامج التالى يقوم بحساب الآتى :

- الربيع الأعلى.
- الربيع الأدنى.
- الوسيط.
- الانحراف الربيعى.
- معامل الالتواء.

وهو هنا يقوم بحساب معامل الالتواء باستخدام طريقة الربيعات التى سبق شرحها فى هذا الفصل باستخدام المعادلة :

$$Q = \frac{(Q_1 - 2Q_2 + Q_3)}{Q_1 - Q_3}$$

حيث :

Q = معامل الالتواء

Q_1 = الربيع الأعلى

Q_2 = الوسيط

Q_3 = الربيع الأدنى

```
10  REM        برنامج لحساب الالتواء بطريقه الربيعـــات
20  DIM A(7),B(7),C(7),D(7),E(7),F(7),G(7)
30  T=0
40  READ N
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)     REM الحد الاعلى، الحد الادنى و التكرار
70   T=T+F(I)                REM مجموع التكرارات
80  NEXT I
90  C(1)=F(1)
100 FOR I=2 TO N
110  C(I)=C(I-1)+F(I)
120 NEXT I
130 :         التجمع التكرارى       التكرار      الفئـــه      رقم
140 :            الصـــاعـد                                     الفئه
150 :        ______________________________________________________
160 :              ###               ##      ##.# - ##.#      ##
170 :                               ___
180 :                               ###                     المجموع
190 PRINT USING 150
195 PRINT USING 130
200 PRINT USING 140
210 PRINT USING 150
220 PRINT
230 FOR I=1 TO N
240  PRINT USING 160, C(I),F(I),B(I),A(I),I
250 NEXT I
260 PRINT USING 170
270 PRINT USING 180 ,T
280 PRINT
290 S1=T*3/4
295 S2=T/2
300 S3=T/4
310 FOR I=1 TO N
320  IF C(I)>S1 THEN 340
330  J=I
340 NEXT I
342 FOR I=1 TO N
344  IF C(I)>S2 THEN 348
346  K=I
348 NEXT I
350 FOR I=1 TO N
360  IF C(I)>S3 THEN 380
370  V=I
380 NEXT I
400 PRINT
450 GOSUB 610
460 Q1=FNA(L1,S1,O1,P1,W1)
470 PRINT ,Q1;' = الربيع الأعلى'
480 PRINT
485 GOSUB 651
490 Q2=FNA(L2,S2,O2,P2,W2)
495 PRINT ,Q2;' = الوسيط'
500 PRINT
510 GOSUB 660
520 Q3=FNA(L3,S3,O3,P3,W3)
530 PRINT ,Q3;' = الربيع الأدنى'
540 PRINT
550 PRINT
560 Y=(Q1-Q3)/2
570 PRINT ,Y;' = الانحراف الربيعى'
572 PRINT
573 PRINT
575 Q=(Q1-2*Q2+Q3)/(Q1-Q3)
578 PRINT USING 580, Q
580 :                        #.#### = معامل الالتواء
585 PRINT
590 PRINT
600 STOP
610 L1=A(J+1)
620 O1=C(J)
630 P1=B(J)-A(J)
640 W1=F(J+1)
650 RETURN
```

```
651 L2=A(K+1)
652 O2=C(K)
653 P2=B(K)-A(K)
654 W2=F(K+1)
655 RETURN
660 L3=A(V+1)
670 O3=C(V)
680 P3=B(V)-A(V)
690 W3=F(V+1)
700 RETURN
710 DEF FNA(L,X,O,P,W)=L+((X-O)*P)/W
720 DATA 7,3,7,1,7,11,3,11,15,11,15,19,20,19,23,9,23,27,4,27,31,2
750 END
```

المخرجات

رقم الفئة	الفئـــه	التكرار	التجمع التكراري الصاعد
1	7.0 - 3.0	1	1
2	11.0 - 7.0	3	4
3	15.0 - 11.0	11	15
4	19.0 - 15.0	20	35
5	23.0 - 19.0	9	44
6	27.0 - 23.0	4	48
7	31.0 - 27.0	2	50
المجموع		50	

الربيع الاعلى = 20.1111

الوسيط = 17

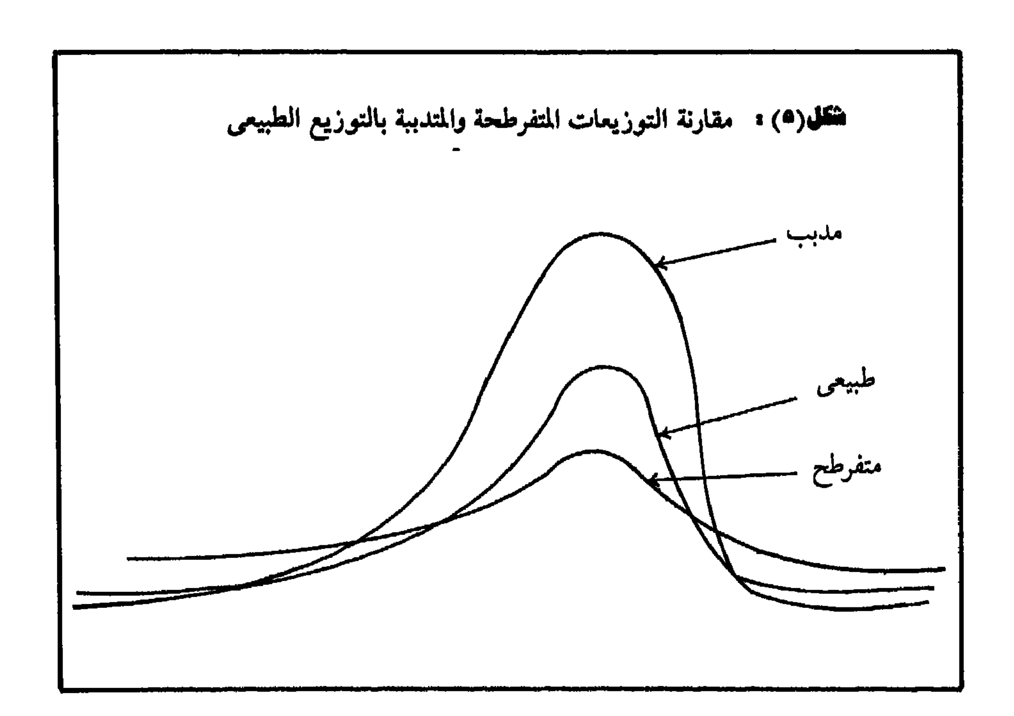
الربيع الادنى = 14.09091

الانحراف الربيعى = 3.010095

معامل الالتواء = .0336

٩ - التفرطح (Kurtosis) :

التفـرطـح عكس التدبب، فالتوزيع المتفرطح هو الذى يكون أقل ارتفاعاً من التوزيع الطبيعى، أما التوزيع المدبب فهو الأكثر ارتفاعاً من التوزيع الطبيعى ؛ إذاً فالتوزيع المتفرطح يتميز بمعامل اختلاف أكبر من الطبيعى لأن قمته أكثر تسطحاً.



شكل(٥) : مقارنة التوزيعات المتفرطحة والمتدببة بالتوزيع الطبيعى

ويعرف معامل التفرطح بأنه نسبة العزم الرابع إلى مربع العزم الثانى . أى أن :

$$ت_{٢} = \text{معامل التفرطح} = \frac{ل_{٤}}{(ل_{٢})^{٢}} \qquad (٣١)$$

ومعلوم أن العزم الثانى هو التباين، لذلك يزداد المقام كلما ازداد التشتت، أى كلما قل معامل الاختلاف لنفس المفردات. وهذا يعنى أن معامل التفرطح يتناقص كلما ازداد التشتت، أى كلما ازداد التفرطح . ومعامل التفرطح للتوزيع الطبيعى الأمثل يساوى ثلاثة . إذاً فالتوزيع مائل نحو التدبب إذا زاد المعامل على ثلاثة، ومائل نحو التفرطح إذا قل عن ذلك . ويلاحظ أن بسط معامل التفرطح وكذلك مقامه لا يكونان سالبين؛ لذلك تكون قيمة التفرطح موجبة فى جميع الحالات ولا تساوى صفراً، إلا إذا كانت جميع القيم متساوية . أما العزوم الفردية وهى $ل_{١}$ ، $ل_{٣}$ ، $ل_{٥}$ ... فتساوى صفراً إذا كان التوزيع متماثلاً كما ورد من قبل .

مثال (١٠,٥) :

أوجد معامل التفرطح للبيانات أدناه :

الفئات	ك ر	س ر - س̄	(س ر - س̄)^{2} ك ر	(س ر - س̄)^{4} ك ر
٣ - ٧	١	-١٢,٢٤ *	١٤٩,٨١٧٦	٢٢٤٤٥,٣١٣٢٧
٧ - ١١	٣	-٨,٢٤	٢٠٣,٦٩٢٨	١٣٨٣٠,٢٥٢٢٦
١١ - ١٥	١١	-٤,٢٤	١٩٧,٧٥٣٦	٣٥٥٥,١٣٥١١٩
١٥ - ١٩	٢٠	-٠,٢٤	١,١٥٢	٠,٠٦٦٣٥٥٢
١٩ - ٢٣	٩	٣,٣٦	١٢٧,٢٣٨٤	١٧٩٨,٨٦٤
٢٣ - ٢٧	٤	٧,٧٦	٢٤٠,٨٧٠٤	١٤٥٠٤,٦٣٧٤
٢٧ - ٣١	٢	١١,٧٦	٢٧٦,٥٩٥٢	٣٨٢٥٢,٤٥٢٣٣
المجموع	٥٠		١١٩٧,١٢٠	٩٤٣٨٦,٦٩

* س̄ = ١٧,٢٤

$$\text{ل}_{2} = \frac{١١٩٧,١٢٠}{٥٠}$$

$$= ٢٣,٩٤٢$$

$$\text{ل}_{4} = \frac{٩٤٣٨٦,٦٩}{٥٠} = ١٨٨٧,٧٣٤$$

$$\text{ت}_{2} = \text{معامل التفرطح} = \frac{\text{ل}_{4}}{(\text{ل}_{2})^{2}} = ٣,٢٩٣١$$

إذاً فالتوزيع قريب جداً من التماثل، إلا أنه مرتفع قليلاً نحو التدبب، ويمكن التأكد من صحة ذلك بمقارنة قيمة ت_{2} = ٣,٢٧٠ بالقيمة (م_{2}) المجدولة بنهاية الفصل، والتي تساوي ٤,٨٨ . وبما أن القيمة المحسوبة أقل من المجدولة فالنتيجة النهائية هي أن التوزيع لا يختلف اختلافاً جوهرياً عن التوزيع الطبيعي الأمثل.

البرنامج التالى يقوم بحساب الآتى :

- العزم الثانى .
- العزم الثالث.
- العزم الرابع .
- التفرطح.
- الالتواء.
- الوسط الحسابى .
- معامل التفرطح .

وباستخدام المعادلات :

$$M_2 = \frac{P}{T} \quad ; \quad M_4 = \frac{R}{T} \quad ; \quad K_1 = \frac{M_4}{(M^2)_2} \qquad S_1 = \frac{M_4}{(M^2)_3}$$

حيث :

M_2	=	العزم الثانى
M_4	=	العزم الرابع
K_1	=	التفرطح
S_1	=	الالتواء
P	=	مجموع مربعات الانحرافات مرجحة بالتكرارات
T	=	مجموع التكرارات

```
10  REM          برنامج لحساب التفرطح
20  DIM A(7),B(7),C(7),D(7),E(7),F(7),H(7),L(7)
22  P=0
23  Q=0
24  R=0
25  S=0
30  T=0
35  D1=0
40  READ N
50  FOR I=1 TO N
60   READ A(I),B(I),F(I)
70   T=T+F(I)
80   C(I)=(A(I)+B(I))/2
90   D(I)=C(I)*F(I)
100 D1=D1+D(I)
110 NEXT I
120 M=D1/T
130 FOR I=1 TO N
140  E(I)=C(I)-M
150  L(I)=(C(I)-M)**2*F(I)
170  H(I)=(C(I)-M)**4*F(I)
180  P=P+L(I)
190  R=R+H(I)
200 NEXT I
210 PRINT USING 250
220 PRINT USING 240
230 PRINT USING 250
240 :     (س-س_)٤ك        (س ر-س_)٢ك      (س ر-س_)      ك        الحـــد
250 :  ____________________________________________________________________
260 FOR I=1 TO N
270  PRINT
280  PRINT USING 290, H(I),L(I),E(I),F(I),B(I),A(I)
290 :     ######.###      ####.###      ###.###      ###     ### - ###
300 NEXT I
310 PRINT USING 320
320 :    __________      ________                   ___
330 PRINT
340 PRINT USING 350, R,P,T
350 :    #######.###     #####.###                  ###      المجموع
360 M2=P/T         REM SECOND MOMENT
370 M4=R/T         REM FOURTH MOMENT
380 K1=M4/M2**2    REM KURTOSIS
400 S1=M4/M2**3    REM SKEWNESS
410 PRINT
420 PRINT ,M2;'= العزم الثاني'
430 PRINT ,Q1;'= العزم الثالث'
440 PRINT ,M4;'= العزم الرابع'
450 PRINT ,K1;'= التفرطح'
460 PRINT ,S1;'= الالتواء'
470 PRINT
480 PRINT
490 PRINT ,M;'الوسط الحسابي='
500 PRINT
510 PRINT ,K1;'معامل التفرطح='
520 PRINT
530 DATA 7,3,7,1,7,11,3,11,15,11,15,19,20,19,23,9,23,27,4,27,31,2
540 END
```

المخرجات

(س-س_)٤ك	(س ز-س_)٢ك	(س ر-س_)	ك	الحـــد
22445.240	149.817	-12.240	1	7 - 3
13830.180	203.692	-8.240	3	11 - 7
3555.097	197.753	-4.240	11	15 - 11
0.066	1.152	-0.240	20	19' - 15
1798.864	127.239	3.760	9	23 - 19
14504.700	240.871	7.760	4	27 - 23
38252.570	276.595	11.760	2	31 - 27
94386.690	1197.119		50	المجموع

```
23.94238  = العزم الثاني
0         = العزم الثالث
1887.734  = العزم الرابع
3.293108  = التفرطح
.137543   = الالتواء

17.23999  الوسط الحسابي=

3.293108  معامل التفرطح=
```

البرنامج التالى يقوم بحساب الآتى :

- العزم الثانى .
- العزم الثالث.
- العزم الرابع .
- التفرطح .
- الالتواء.

وتستخدم فيه نفس البيانات المستخدمة فى المثال (٥ , ٥).

والمعادلات :

$$R_1 = R/T \quad ; Q_1 = Q/T \; ; \; P_1 = P/T \; ;$$

$$K_1 = R_1/(P_1)^2 \; ; \; S_1 = (Q_1)^2 /(P_1)^3$$

حيث :

R_1 = العزم الرابع
Q_1 = العزم الثالث
P_1 = العزم الثانى
T = مجموع التكرارات
K_1 = التفرطح
S_1 = الالتواء

```
10 REM                         برنامج لحساب العزوم
20 DIM A(13),B(13),C(13),D(13),E(13),F(13),G(13),H(13),K(13),L(13)
22 P=0
23 Q=0
24 R=0
25 S=0
30 T=0
35 D1=0
40 READ N   REM  عدد الفئات
50 FOR I=1 TO N
60 READ A(I),B(I),F(I)  REM الحد الادنى، الحد الاعلى، التكرار
70 T=T+F(I)
80 C(I)=(A(I)+B(I))/2
90 D(I)=C(I)*F(I)
100 D1=D1+D(I)
110 NEXT I
120 M=D1/T
130 FOR I=1 TO N
140 E(I)=C(I)-M
150 L(I)=(C(I)-M)**2*F(I)
160 G(I)=(C(I)-M)**3*F(I)
170 H(I)=(C(I)-M)**4*F(I)
180 K(I)=ABS(E(I))*F(I)
181 P=P+L(I)
182 Q=Q+G(I)
183 R=R+H(I)
184 S=S+K(I)
190 NEXT I
200 PRINT USING 215
205 PRINT USING 210
207 PRINT USING 215
210 : |س-س_|ك  (س-س_)٤ك  (س-س_)٣ك  (س-س_)٢ك  (س-س_)  س  ك  الفئة
215 : ______________________________________________________________
220 FOR I=1 TO N
230 PRINT
240 PRINT USING 250, K(I),H(I),G(I),L(I),E(I),C(I),F(I),B(I),A(I)
250 : ###.###  ######.###  ######.###  ####.###  ###.###  ###   ###  ##.# - ##.=
260 NEXT I
265 PRINT USING 270
266 PRINT USING 275, S,R,Q,P,T
270 : ________  __________  __________  __________               _____
275 : ###.####  ######.##  #####.####  #####.###               ###     المجموع
290 PRINT
300 R1=R/T        REM FOURTH MOMENT
310 Q1=Q/T        REM THIRD MOMENT
320 P1=P/T        REM SECOND MOMENT
330 K1=R1/P1**2     REM KURTOSIS
340 S1=(Q1/P1**(3/2))**2    REM SKEWNESS
350 PRINT
360 PRINT ,R1;'= العزم الرابع'
370 PRINT ,Q1;'= العزم الثالث'
380 PRINT ,P1;'= العزم الثاني'
390 PRINT ,K1;'= التفرطح'
400 PRINT ,S1;'= الالتواء'
410 PRINT
420 PRINT
430 PRINT ,M;'=الوسط الحسابي'
440 PRINT
450 PRINT
460 DATA 13,12.5,15.5,2,15.5,18.5,3,18.5,21.5,4,21.5,24.5,5
470 DATA 24.5,27.5,14,27.5,30.5,21,30.5,33.5,59,33.5,36.5,24
480 DATA 36.5,39.5,7,39.5,42.5,6,42.5,45.5,2,45.5,48.5,1,48.5,51.5,2
490 END
```

المخرجات

|س-س_|ك	(س-س_)٤ك	(س-س_)٣ك	(س-س_)٢ك	(س-س_)	س	ك	الفئة
34.960	186721.600	-10682.020	611.100	-17.480	14	2	15.5 - 12.5
43.440	131884.800	-9108.070	629.011	-14.480	17	3	18.5 - 15.5
45.920	69474.690	-6051.805	527.161	-11.480	20	4	21.5 - 18.5
42.400	25855.470	-3048.995	339.552	-8.480	23	5	24.5 - 21.5
76.720	12625.500	-2303.926	420.425	-5.480	26	14	27.5 - 24.5
52.080	794.370	- - -320.311	129.158	-2.480	29	21	30.5 - 27.5
30.680	4.314	8.296	15.954	0.520	32	59	33.5 - 30.5
84.480	3684.546	1046.745	297.370	3.520	35	24	36.5 - 33.5
45.640	12649.970	1940.177	297.573	6.520	38	7	39.5 - 36.5
57.120	49283.300	5176.812	543.783	9.520	41	6	42.5 - 39.5
25.040	49141.440	3925.034	311.501	12.520	44	2	45.5 - 42.5
15.520	58018.610	3738.312	240.871	15.520	47	1	48.5 - 45.5
37.040	235284.900	12704.370	685.981	18.520	50	2	51.5 - 48.5
591.0391	835423.20	-2975.3750	5071.434			150	المجموع

العزم الرابع = 5569.488
العزم الثالث = -19.83593
العزم الثاني = 33.80956
التفرطح = 4.872326
الالتواء = 1.018082E-02

الوسط الحسابي= 31.48

جدول (م ١) الالتواء

حجم العينة	النسبة المئوية 5%	1%	الانحراف المعيارى	حجم العينة	النسبة المئوية 5%	1%	الانحراف المعيارى
25	0.711	1.061	0.4354	100	0.389	0.567	0.2377
30	0.662	0.986	0.4052	125	0.350	0.508	.2139
35	0.621	0.923	0.3804	150	0.321	0.464	.1961
40	0.587	0.870	0.3596	175	0.298	0.430	.1820
45	0.558	0.825	0.3418	200	0.280	0.403	.1706
50	0.534	0.787	.3264	250	0.251	0.360	.1531
60	0.492	0.723	0.3009	300	0.230	0.329	.1400
70	0.459	0.673	0.2806	350	0.213	0.305	.1298
80	0.432	0.631	0.2638	400	0.200	0.285	.1216
90	0.409	0.596	0.2498	450	0.188	0.269	.1147
100	0.389	0.567	0.2377	500	0.179	0.255	.1089

جدول (م ٢) التفرطح

حجم العينة	النسبة المئوية أعلى 5%	أعلى 1%	أدنى 5%	أدنى 1%	حجم العينة	النسبة المئوية أعلى 5%	أعلى 1%	أدنى 5%	أدنى 1%
50	4.88	3.99	2.15	1.95	600	3.54	3.34	02.70	2.60
75	4.59	3.87	2.27	2.08	650	3.52	3.33	02.71	2.61
100	4.39	3.77	2.35	2.18	700	3.50	3.31	02.72	2.62
125	4.24	3.71	2.40	2.24	750	3.48	3.30	02.73	2.64
150	4.13	3.65	2.45	2.29	800	3.46	3.29	02.74	2.65
					850	3.45	3.28	02.74	2.66
200	3.98	3.57	2.51	2.37	900	3.43	3.28	02.75	2.66
250	3.87	3.52	2.55	2.42	950	3.42	3.27	02.76	2.67
300	3.79	3.47	2.39	2.46	1000	3.41	3.26	02.76	2.68
350	3.72	3.44	2.62	2.50					
400	3.67	3.41	2.64	2.52	1200	3.37	3.24	02.78	2.71
450	3.63	3.39	2.66	2.55	1400	3.34	3.22	02.80	2.72
500	3.60	3.37	2.67	2.57	1600	3.32	3.21	02.81	2.74
550	3.57	3.35	2.69	2.58	1800	3.30	3.20	02.82	2.76
600	3.54	3.34	2.70	2.60	2000	3.28	3.18	02.83	2.77

المصدر : Snedcor (G.W.) and Cochran (W.G.); Statistical Methods; Jwo P. Press; IWOA, 1974; Table (6).

تمارين

(١) قيست خمس قطع بأربع طرق مختلفة وكانت النتائج على النحو الآتى :

القطعة	الطريقة (أ)	الطريقة (ب)	الطريقة (جـ)	الطريقة (د)
١	٣	٠,٠٣	١٠٣	١٠٣٠
٢	٦	٠,٠٦	١٠٦	١٠٦٠
٣	٩	٠,٠٩	١٠٩	١٠٩٠
٤	٥	٠,٠٥	١٠٥	١٠٥٠
٥	٧	٠,٠٧	١٠٧	١٠٧٠

فأوجد الانحراف المعيارى للطول بكل طريقة.

(٢) إذا كانت س عينة من عدة متغيرات قوامها (ن) وحدة، وإذا كان الانحراف المعيارى لتلك العينة ع، فاستخدم نتائج السؤال الأول لإيجاد الانحراف المعيارى للعينة ص المكونة من (ن) وحدة إذا كان :

(أ) ص = س + ل

(ب) ص = ل س

علماً بأن ل مقدار (عدد) ثابت.

(٣) استخدم بيانات السؤال الأول لإيجاد المدى لكل طريقة وقارن بين النتائج.

(٤) أى الطرق الواردة فى السؤال الأول كانت نتائجها أكثر تجانساً؟

(٥) أوجد الانحراف المعيارى للبيانات التالية الخاصة بتوزيع بعض المصابين فى حوادث المرور حسب الأعمار، علماً بأن البيانات كانت خلال أحد مواسم الحج، والبيانات هى :

العمر بالسنوات	عدد المصابين فى الرياض	عدد المصابين فى جدة
١ - ٦	١	١
٧ - ١٢	٧	٣
١٣ - ١٨	١٥	٧
١٩ - ٢١	٣٠	١٠
٢١ - ٢٥	٢٧	١١
٢٥ - ٣٠	١٢	١٤
٣١ - ٣٨	٥	١٦
٣٩ - ٤٨	٢	١٨
٤٩ - ٥٨	١	٢٠

(٦) أى المدينتين فى السؤال الخامس أكثر تشتتاً؟

(٧) أوجد التواء وتفرطح البيانات الخاصة بكل مدينة فى السؤال الخامس، وهل يمكن اعتبار التوزيعين طبيعيين؟

(٨) استخدم بيانات السؤال الخامس، لإيجاد الانحراف الربيعى لبيانات كل مدينة.

(٩) افتتحت شركة سعودية فرعاً لها فى السودان، فإذا كان الوسط الحسابى لرواتب موظفيها فى الرياض ٦٠٠٠ ريال، بينما كان الوسط الحسابى لرواتب موظفيها فى الخرطوم ٦٠٠ جنيه، بينما كان التباين فى الرياض ١٦٠٠٠٠ (ريال)^2 وفى الخرطوم ٩٠٠ (جنيه)^2، فأوجد أى الموظفين أفضل مقارنة بزملائه فى كل حالة :

(أ) موظف راتبه فى الرياض ٧٠٠٠ريال، وموظف راتبه فى الخرطوم ٧٠٠جنيه.
(ب) موظف راتبه فى الرياض ٦٠٠٠ريال، وموظف راتبه فى الخرطوم ٦٠٠جنيه.
(جـ) موظف راتبه فى الرياض ٥٠٠٠ريال، وموظف راتبه فى الخرطوم ٥٠٠جنيه.

(١٠) البيانـات التـالية تمثل تغيرات أسعار نفس النوع من التفاح خلال نفس الفترات فى مدينتين، والأسعار هى :

الفترة	السعر فى المدينة (أ) للكيلو الواحد بالريال	السعر فى المدينة (ب) للكيلو الواحد بالريال
١	٢٥	٢٣
٢	٢٨	٢٤
٣	٢٢	٢٤
٤	٢٦	٢٦
٥	٣٥	٣٦
٦	٢٠	١٦

فأوجد التباين للأسعار فى كل مدينة .

(١١) أى المدينتين أكثر استقراراً فى سعر التفاح؟

(١٢) أوجد التباين للمدينتين معاً؟ هل يساوى مجموع التباينين؟ ولماذا؟

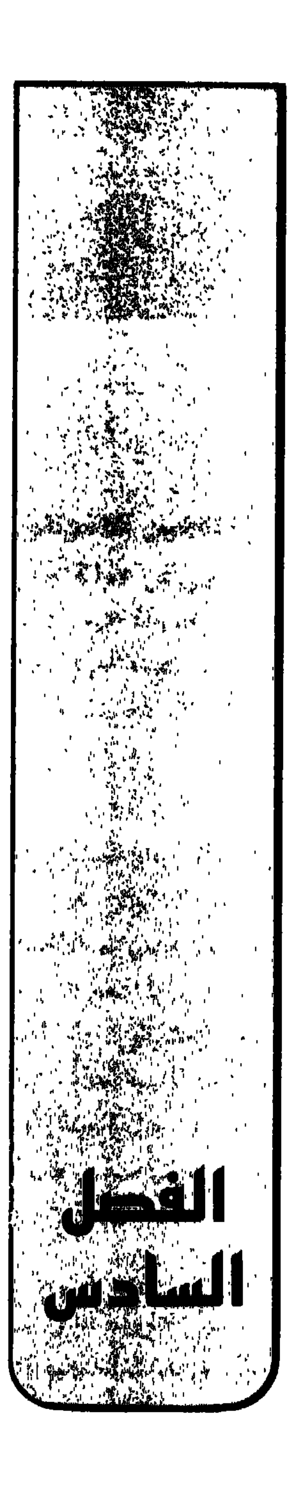
(١٣) اكتب برنامج بيسك لإيجاد الانحراف المعيارى للبيانات الواردة بالسؤال (١) .

(١٤) اكتب برنامج بيسك لإيجاد الانحراف المعيارى للبيانات الواردة بالسؤال (٥) .

(١٥) اكتب برنامج بيسك لإيجاد قيمة الالتواء والتفرطح لكل مدينة حسب البيانات فى السؤال (٥)

(١٦) اكتب برنامج بيسك لإيجاد التباين للبيانات الواردة بالسؤال (١٠) .





أهم التوزيعات الاحتمالية





أهم التوزيعات الاحتمالية

١- المتغير العشوائي (Random Variable) :

تسمى كل عملية أجريت ولا يمكن التأكد من نتيجتها مسبقاً بالتجربة الإحصائية، وتسمى مجموعة جميع النتائج المتوقعة بفضاء العينة (Sample Space) ، كما تسمى أى مجموعة جزئية من فضاء العينة بالحادث (Event) . فرمى زهر الطاولة (النرد) تجربة إحصائية، والنقاط ١، ٢، ٣، ٤، ٥، ٦، تمثل فضاء العينة، وأى نقطة من هذه النقاط أو مجموعة نقاط تمثل حادثاً.

أما الاختيار العشوائى فهو الذى استخدمت فيه إحدى طرق الاختيار العشوائى، وطرق الاختيار العشوائى هى التى تعطى احتمالاً متساوياً لاختيار أى عنصر من العناصر المكونة للمجتمع الذى تم الاختيار منه. إذاً فاختيار أى نقطة من نقاط الزهر عند رميه يعتبر اختياراً عشوائياً. وأما المتغير الذى يمكن أن يظهر فى مدى - حيز - معين، وباحتمالات مختلفة، فيسمى بالمتغير العشوائى. وأما الحيز الذى تعرف به جميع المتغيرات العشوائية الخاصة بتجربة معينة، فهو مجال للعينة أيضاً. لذلك فالمتغير العشوائى هو دالة ذات قيمة عددية (سالبة أو موجبة) معرفة على مجال العينة. فمجموع أى نقطتين مثلاً عند رمى زهر طاولة مرة واحدة يمثل متغيراً عشوائياً ضمن مجال العينة (فضاء أو فراغ) ٢، ٣، ٤، ٥، ٦، ٧، ٨، ٩، ١٠، ١١، ١٢.

يسمى التوزيع الذى يمثل الاحتمالات الخاصة المتغيرات العشوائية بدالة توزيع الاحتمال، وهناك توزيعات احتمالية كثيرة سيعرض هنا أكثرها استخداماً.

٢ ـ التوزيع الطبيعى (Normal Distribution) :

هو أكثر التوزيعات الاحتمالية استخداماً؛ لأن توزيع المتغيرات يكون طبيعياً فى أكثر الحالات التطبيقية، كما أنه يمثل تقديراً دقيقاً لعدد كبير من التوزيعات الأخرى إذا كان عدد المتغيرات (ن) كبيراً.

لكل توزيع معالم، وللتوزيع الطبيعى مَعْلَمان (Parameters) هما: الوسط الحسابى للمجتمع (و) والانحراف المعيارى لذلك المجتمع (م). وإذا كانت ط تمثل القيمة الثابتة ٣,١٤ بينما تمثل هـ القيمة الأسية (Expoent) التى تساوى ٢,٧١٨ تقريباً فإن دالة توزيع المتغير العشوائى (ق(س)) هى :

$$ق(س) = \frac{١}{م\sqrt{٢ ط}} هـ^{-(س - و)^٢/٢م^٢} \qquad (١)$$

$$وتسمى القيمة ى = \frac{(س - و)}{م} \quad بالقيمة المعيارية$$

لذلك عُرِّف التوزيع الطبيعى المعيارى (Standard Normal Dist.) بأنه توزيع طبيعى بوسط حسابى قدره صفر، وانحراف معيارى قدره واحد، ودالة توزيع على النحو التالى :

$$ق(ى) = \frac{١}{\sqrt{٢ط}} هـ^{-\frac{ى^٢}{٢}} \qquad (٢)$$

ولكثرة استخدامات التوزيع الطبيعى المعيارى فقد دون فى جداول (ملحق رقم ١)؛ ليسهل استخدامه فى الحالات التطبيقية. هذا ويلاحظ أن حدود الجدول هى - ٣ إلى + ٣ وذلك لأن ٩٩,٧٪ من الحالات تنحصر فى المساحة الواقعة بين - ٣ و + ٣ للتوزيع الطبيعى المعيارى (راجع الانحراف المعيارى والمقارنات بالفصل السابق).

أما المساحة الكلية لمنحنى دالة التوزيع الطبيعى المعيارى فتساوى الوحدة، نصفها فوق الصفر، ونصفها تحته (٠,٥)؛ لذلك فإن جميع القيم المعيارية الموضحة على الجدول تزيد على ٠,٥ إذا كانت أكبر من الصفر، وتقل عنها إذا كانت أقل من الصفر. فهى إذاً توضح النسبة التى تحت القيمة المعيارية المعينة.

مثال (٦,١) :

أوجـد نسبة الطلاب الذين تزيد درجاتهم على ٨٢ درجة فى مادة الرياضيات، إذا كان الوسط الحسابى للممتحنين ٧٦ درجة بانحراف معيارى قدره ٤ .

الحل :

$$ق(ى) = ق\left(\frac{س_ر - \bar{س}}{ع}\right) \quad (٣)$$

$$= ق\left(\frac{٨٢ - ٧٦}{٤}\right)$$

$$= ق\left(\frac{٦}{٤}\right)$$

$$= ق\,(١,٥)$$

القيمة المناظرة بالجدول لقيمة ق (١,٥) هى ٠,٩٣٣

إذاً هناك ٩٣,٣٪ من الطلاب لا تزيد درجاتهم على ٨٢، وعليه فنسبة الذين يزيدون على ٨٢ درجة هى :

$$١٠٠ - ٩٣,٣ = ٦,٧٪$$

هذا وتمثل المساحة الواقعة تحت المنحنى الاحتمال أيضاً؛ لذلك فإن احتمال اختيار طالب واحد عشوائياً تكون درجته ٨٢ فما فوق يساوى ٠,٠٦٧

مثال (٦,٢) :

المسـافـة التى يقطعها أحد العدائين ذات توزيع طبيعى بوسط حسابى قدره ٣ دقائق، وانحـراف معيارى يسـاوى ٢٠ ثانية، فما احتمال أن يقطع نفس المسافة فى مدة تتراوح بين دقيقتين وخمسين ثانية وثلاث دقائق وعشرين ثانية؟

الحل :

$\bar{س}$ = ١٨٠ ثانية

ع = ٢٠ ثانية

الاحتمال (ح) هو المساحة المحصورة بين ٢٠٠ ثانية و ١٧٠ ثانية، وعليه تكون :

$$ق (ى) \quad = \quad ق \left(\frac{٢٠٠ - ١٨٠}{٢٠}\right) \quad - \quad ق \left(\frac{١٧٠ - ١٨٠}{٢٠}\right)$$

$$= \quad ق(١) - ق(-\frac{١}{٢})$$

وباستخدام الجدول :

$$ح \quad = \quad ق (ى) \quad = \quad ٠,٨٤١٣ - ٠,٣٠٨٥$$

$$ح \quad = \quad ٠,٥٣٢٨$$

٣ - التوزيع ذو الحدين (Binomial Distribution) :

هو توزيع كثير الاستخدام فى حالة المتغيرات التى تنقسم إلى قسمين فقط، أحدهما يحمل صفة معينة والآخر لا يتمتع بتلك الصفة، فأوجه زهرة النرد بعضها يتميز بالأعداد الفردية وبعضها بالزوجية، والمواليد ذكور أو إناث، والممتحنون بعضهم ناجح وبعضهم فاشل، وكذلك المجيبون عن الأسئلة التى تحدد إجاباتها بنعم أو لا. ويشترط أن تكون كل تجربة مستقلة عن الأخرى، حتى تتبع المتغيرات العشوائية التوزيع ذا الحدين الذى يتميز بمعلمين، هما : عدد القيم العينية (ن)، واحتمال أن تنتهى التجربة بنجاح (جـ)، وكلمة نجاح هنا تعنى حدوث الحدث المعين سواء كان ذلك الحدث مرغوباً فيه أم لا.

فإذا كانت (س) تمثل عدد حالات النجاح عند تكرار تجربة ذى الحدين لعدد ن مرة فمعادلة التوزيع ذى الحدين هى :

$$ح (س) \quad = \quad \binom{ن}{س} \quad (جـ)^{س} \quad (١ - جـ)^{ن - س} \qquad (٤)$$

$$س \quad = \quad ٠، ١، ٢، ٣، ....، ن$$

و ح (س) تعنى احتمال النجاح س مرة.

مثال (٣,٦) :

احتمال أن تكون الإجابة بنعم عن أحد الأسئلة هو $\frac{٣}{٤}$ ، فإذا أخذت عينة عشوائية من ٥ إجابات، فما احتمال أن تكون :

أ - كل الإجابات بنعم؟

ب - كلها لا؟

جـ - ٣ حالات بنعم؟

د - ٣ حالات على الأقل بنعم؟

الحل :

يمكن كتابة $\binom{ن}{س}$ على هذا النحو : $\frac{ن\,!}{س\,!\,(ن - س)\,!}$

حيث ن ! هو مضروب ن وهو ضرب كل الأعداد الموجبة من ١ إلى ن فمثلاً

$$٦\,! = ٦×٥×٤×٣×٢×١ .$$

هذا وتسمى $\binom{ن}{س}$ بالتوافيق، وهى عدد المجموعات الجزئية التى تتكون كل مجموعة منها من (س) عنصر، والتى يمكن استخراجها من مجموعة مكونة من ن عنصراً.

هذا وتجدر الإشارة إلى أن ١ ! = ٠ ! = ١، أى أن مضروب الواحد يساوى مضروب الصفر.

وبتطبيق ذلك على السؤال السابق يمكن استخراج الاحتمال فى كل حالة كما يلى، علماً بأن التوافيق يمكن استخراجها بواسطة تطبيق المعادلة السابقة أو الآلات الحاسبة العلمية أو الحاسبات الآلية أو استخدام الجداول (ملحق رقم ٥).

(أ) $ح(٥) = \binom{٥}{٥}\left(\frac{٣}{٤}\right)^{٥}\left(\frac{١}{٤}\right)^{٠}$

$$= \left(\frac{٣}{٤}\right)^{٥} = ٠,٢٣٧$$

(ب) $$\text{ح}(٠) = \binom{٥}{٠}\left(\frac{٣}{٤}\right)^{٠}\left(\frac{١}{٤}\right)^{٥}$$

$$= \left(\frac{١}{٤}\right)^{٥}$$

$$= ٠,٠٠١$$

(ج) $$\text{ح}(٣) = \binom{٥}{٣}\left(\frac{٣}{٤}\right)^{٣}\left(\frac{١}{٤}\right)^{٢}$$

$$= ١٠ \times \frac{٢٧}{٦٤} \times \frac{١}{١٦}$$

$$= ٠,٢٦٤$$

(د) $$\text{ح}(\text{س} \geqslant ٣) = \text{ح}(٣) + \text{ح}(٤) + \text{ح}(٥)$$

$$= \binom{٥}{٣}\left(\frac{٣}{٤}\right)^{٣}\left(\frac{١}{٤}\right)^{٢} + \binom{٥}{٤}\left(\frac{٣}{٤}\right)^{٤}\left(\frac{١}{٤}\right)^{١} + \binom{٥}{٥}\left(\frac{٣}{٤}\right)^{٥}\left(\frac{١}{٤}\right)^{٠}$$

$$= \sum_{\text{س}=٣}^{٥} \binom{\text{ن}}{\text{س}} (\text{جـ})^{\text{س}} (١ - \text{جـ})^{\text{ن} - \text{س}}$$

$$= ٠,٢٦٤ + ٠,٣٩٦ + ٠,٢٣٧$$

$$= ٠,٨٩٧$$

أما الوسط للمتغيرات التى تتبع التوزيع ذا الحدين فهو ن جـ وأما الانحراف المعيارى فهو $\sqrt{\text{ن جـ } (١ - \text{جـ})}$، أى أن التباين هو ن جـ (١ ـ جـ). فعدد الإجابات المتوقع أن تكون بنعم من بين الإجابات السابقة هو :

$$٥ \times \frac{٣}{٤} = ٣,٧٥$$

والانحراف المعيارى لذلك التقدير هو :

$$\sqrt{٥ \times \frac{٣}{٤} \times \frac{١}{٤}} = ٠,٩٦٨$$

يعتبر التوزيع ذو الحدين هو أقدم التوزيعات، إذ سبق التوزيع الطبيعى بعشرين عاماً. هذا ولقد عرف التوزيع الطبيعى عندما كان أحد العلماء (De Moivre) يحاول إيجاد معادلة للتوزيع ذى الحدين عندما يكون عدد المتغيرات (ن) كبيراً، ولقد فوجىء ذلك العالم بأن التوزيع المنفصل يتحول إلى توزيع متصل هو التوزيع الطبيعى عندما يكون حجم العينة كبيراً. ولقد تبعت ذا الحدين فى ذلك أكثر التوزيعات الاحتمالية، وهو ما عرف أخيراً بنظرية النهاية المركزية (Central Limit Theorem).

٤ - توزيع مربع كاى (X^2 chi square distribution)

اتضح مما مضى أنه إذا كانت س ذات توزيع طبيعى بوسط قدره (و) وانحراف معيارى (م) فإن :

$$ى = \frac{س - و}{م} \qquad (٥)$$

ذات توزيع طبيعى أيضاً بوسط صفر وانحراف معيارى واحد.

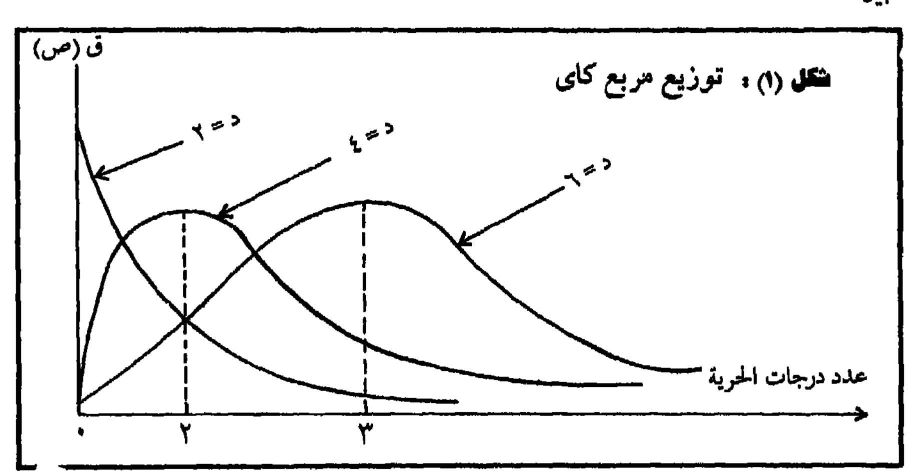
$$أما \quad ص = \frac{(س - و)^٢}{م^٢} \qquad (٦)$$

فهى ذات توزيع جديد يسمى مربع كاى ($ك^٢_د$)، ولمربع كاى معلم واحد يسمى درجات الحرية (د). ومربع التوزيع الطبيعى المعيارى يتبع توزيع مربع كاى بدرجة حرية واحدة ($ك^٢_١$). وأما معادلته فهى على النحو التالى :

$$ق(ص) = \frac{ص^{-\frac{١}{٢}} \; هـ^{-ص/٢}}{\sqrt{٢ ط}} \qquad (٧)$$

حيث ق تعنى الاقتران (الدالة) والرموز ص، ط، هـ كان قد تم تعريفها من قبل.

هذا، ويساوى وسط مربع كاى عدد درجات الحرية (د) بينما يساوى تباينه ضعف درجات الحرية (٢ د). أما شكل التوزيع فيبلغ قمته عند د - ٢ ويقترب من التوزيع الطبيعى كلما ازداد عدد درجات الحرية (د)، إلى أن يصبح توزيعاً طبيعياً، عندما يكون عدد درجات الحرية كبيراً.



شكل (١) : توزيع مربع كاى

يستمد توزيع مربع كاى أهميته فى المجال التطبيقى من الخصائص الثلاث التالية :

الخاصية الأولى :

مجموع المتغيرات العشوائية، التى يتبع كل منها توزيعاً مستقلاً لمربع كاى، يكون تابعاً أيضاً توزيع مربع كاى، وبعدد من درجات الحرية مساوٍ لمجموع درجات حريات التوزيعات المستقلة.

فإذا كان ص = $س_١$ + $س_٢$ حيث $س_١$ تتبع توزيع مربع كاى بعدد من درجات الحرية يساوى $د_١$، و $س_٢$ تتبع توزيعاً مستقلاً لمربع كاى بعدد من درجات الحرية يساوى $د_٢$، فإن ص ذات توزيع لمربع كاى بعدد $د_١$+ $د_٢$ من درجات الحرية.

أما إذا كانت $س_١$، $س_٢$، $س_٣$، . . . ، $س_ن$ عينة عشوائية من توزيع طبيعى بوسط (و)، وانحراف معيارى (م)، وإذا كانت

$$ص = \frac{\sum_{ر=١}^{ن} (س_ر - و)^٢}{م^٢}$$

فإن ص ذات توزيع يتبع لمربع كاى بعدد من درجات الحرية يساوى (ن)؛ ذلك لأن مربع التوزيع الطبيعى المعيارى لكل متغير يتبع توزيع مربع كاى بدرجة حرية واحدة.

الخاصية الثانية :

إذا كانت $س_١$ تتبع توزيع مربع كاى بعدد من درجات الحرية يساوى $د_١$، وإذا كانت $س_٢$ تتبع توزيعاً مستقلاً لمربع كاى بعدد من درجات الحرية يساوى $د_٢$ حيث $د_١ > د_٢$ وإذا كانت $ص = س_١ - س_٢$ فإن ص تتبع توزيع مربع كاى بعدد من درجات الحرية يساوى $د_١ - د_٢$.

الخاصية الثالثة :

إذا كانت $س_١$، $س_٢$، $س_٣$، . . . $س_ن$ عينة عشوائية من توزيع طبيعى بوسط (و)، وانحراف معيارى (م). وإذا كانت $\bar{س}$ و $ع^٢$ هما الوسط الحسابى والتباين الخاصان بالعينة واللذان يستخدمان لتقدير قيمتى (و) و (م) على التوالى، وإذا كانت :

$$ص = \frac{\sum_{ر=١}^{ن} (س_ر - \bar{س})^٢}{م^٢} \qquad (٨)$$

فإن ص تتبع توزيع مربع كاى بعدد من درجات الحرية يساوى (ن - ١)، حيث ن هو حجم العينة.

وبما أن

$$ع^٢ = \frac{\sum (س_ر - \bar{س})^٢}{ن - ١} \qquad (٩)$$

فإن :

$$\frac{(ن - ١) ع^٢}{م^٢} = \frac{\sum (س_ر - \bar{س})^٢}{م^٢} \qquad (١٠)$$

وعليه فإن :

$$\frac{(ن - ١) ع^٢}{م^٢}$$

يتبع أيضاً توزيع مربع كاى بعدد من درجات الحرية يساوى (ن - ١). ولعل هذا هو السبب الذى يجعل قسمة مجموع مربعات الانحرافات على (ن - ١) بدلاً من (ن). هذا، والغريب فى الأمر هو أن سَ و ع ٢ تمثلان إحصائيتين لنفس المتغيرات العشوائية إلا أنهما مستقلتان تماماً بعضهما عن بعض، كذلك يلاحظ أن الفرق بين الخاصية الأولى والثالثة هو استبدال الوسط الخاص بالمجتمع (و) بالوسط الحسابى للعينة (سَ). ولقد أدى هذا الاستبدال إلى نقصان عدد درجات الحرية بواحد؛ وذلك لأن :

$$\frac{\sum (\text{س}_{\text{ر}} - \text{و})^2}{\text{م}^2} = \frac{\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^2}{\text{م}^2} + \left(\frac{\bar{\text{س}} - \text{و}}{\text{م} / \sqrt{\text{ن}}}\right)^2 \quad (١١)$$

والجزء الثانى من الجانب الأيسر فى المعادلة السابقة يتبع توزيع مربع كاى بدرجة حرية واحدة، ويمكننا الآن التأكد من صحة المعادلة الهامة، فالجانب الأيسر من المعادلة هو :

$$\frac{\sum (\text{س}_{\text{ر}} - \bar{\text{س}})^2}{\text{م}^2} + \frac{(\bar{\text{س}} - \text{و})^2}{\text{م}^2} \times \text{ن}$$

يمكن إعادة صياغة البسط على النحو التالى :

$$\sum \text{س}_{\text{ر}}^2 - 2\sum \text{س}_{\text{ر}}\bar{\text{س}} + \text{ن}\,\bar{\text{س}}^2 + \text{ن}\,\bar{\text{س}}^2 - 2\,\text{ن}\,\bar{\text{س}}\,\text{و} + \text{ن}\,\text{و}^2$$

$$= \sum \text{س}_{\text{ر}}^2 - 2\frac{\sum \text{س}_{\text{ر}} \sum \text{س}_{\text{ر}}}{\text{ن}} + \text{ن}\left(\frac{\sum \text{س}_{\text{ر}}}{\text{ن}}\right)^2 + \text{ن}\left(\frac{\sum \text{س}_{\text{ر}}}{\text{ن}}\right)^2 - 2\,\text{ن}\frac{\sum \text{س}_{\text{ر}}\text{و}}{\text{ن}} + \text{ن}\,\text{و}^2$$

$$= \sum \text{س}_{\text{ر}}^2 - 2\frac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}} + \frac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}} + \frac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}} - 2\sum \text{س}_{\text{ر}}\text{و} + \text{ن}\,\text{و}^2$$

$$= \sum \text{س}_{\text{ر}}^2 - 2\sum \text{س}_{\text{ر}}\text{و} + \text{ن}\,\text{و}^2$$

$$= \sum (\text{س}_{\text{ر}} - \text{و})^2 \quad (١٢)$$

وباستخدام الخاصية الثانية للتوزيع يتضح أن استبدال وسط المجتمع بالوسط الحسابى للعينة ينقص عدد درجات الحرية بواحد.

تستخدم جداول مربع كاى (ملحق رقم ٣) لإيجاد القيم التى تقع إلى يسارها أو يمينها مساحة معينة (احتمال).

والمساحة الواقعة إلى يسار القيمة تعنى احتمال (ح) أن تكون قيمة المتغير العشوائى (س) أقل من، أو تساوى، قيمة محددة (ص). وتكتب على النحو التالى :

ح (س ⩽ ص)

لذلك فالمساحة الواقعة إلى يمين قيمة متغير ما، زائداً المساحة إلى يسار نفس قيمة المتغير تساوى واحداً. أى أن :

ح (س ⩽ ص) + ح (س > ص) = ١

هذا، ويلاحظ أن جدول توزيع مربـع كاى قد قسم أفقياً مئينيات معينة ذات دلالات خاصة فى التقدير الإحصائى، واختبار الفروض كما سيتضح فيها بعد، كما أنه قد قسم رأسياً على عدد درجات الحرية التى يقع عليها التوزيع. ولعله من المتوقع أن يزداد الاحتمال أفقياً لنفس العدد من درجات الحرية، لأن قيمة المتغير تزداد أفقياً. أما إذا زاد عدد درجات الحرية، وأردنا الحصول على نفس المساحة الواقعة على يسار (الاحتمال)، فيلزم ذلك زيادة قيمة توزيع مربع كاى.

مثال (٦,٤) :

أوجـد احتـمال أن تكون قيمة المتغير العشوائى (س) التابع لتوزيع مربع كاى على ١٠ درجات حرية أقل من، أو تساوى :

(أ) ٢,١٦

(ب) ١٥,٩٩

(جـ) ٢٣,٢١

الحل :

باستخدام جدول مربع كاى وبالنظر لصف د = ١٠ يتضح أن :

(أ) ح (س ⩽ ٢,١٦) = ١ - ٠,٩٩٥

= ٠,٠٠٥

(ب) ح (س ⩾ ١٥,٩٩) = ١ - ٠,١
= ٠,٩٠

(جـ) ح (س ⩾ ٢٣,٢١) = ١ - ٠,٠١
= ٠,٩٩

مثال (٥ , ٦) :

أوجد قيمة مربع كاى (ص) إذا كان :

(أ) ح (س ⩾ ص) = ٠,٠٠٥

(ب) ح (س ⩾ ص) = ٠,٩٠

(جـ) ح (س ⩾ ص) = ٠,٩٩

علماً بأن المتغير العشوائى (ص) يتبع لتوزيع مربع كاى على ٢٠ درجة حرية .

الحل :

باستخدام جدول مربع كاى وعند تقاطع د = ٢٠ مع الاحتمالات السابقة، يتضح أن :

(أ) ص = ٧,٤٣

(ب) ص = ٢٨,٤

(جـ) ص = ٣٧,٦

هناك علاقة تقريبية بين قيم توزيع مربع كاى والتوزيع الطبيعى المعيارى، عندما يكون عدد درجات الحرية أكبر من ٣٠ . فإذا كانت (أ) تعنى الاحتمال، بمعنى أن

ح (س ⩾ ك) = أ

هى احتمال أن تكون قيمة س أقل من، أو تساوى، قيمة معينة، تساوى (أ) .

وإذا كانت ق (ى) = أ تعني احتمال أن تكون القيمة المعيارية (ى) التابعة للتوزيع الطبيعى المعيارى يساوى (أ) أيضاً فالعلاقة هى :

$$\sqrt{\text{٢ ك}} - \sqrt{\text{٢ د - ١}} = \text{ق (ى)} \qquad (١٣)$$

أى أن :

$$\text{ك} = \frac{1}{2}\left(\text{ق (ى)} + \sqrt{\text{٢ د - ١}}\right)^{2} \qquad (١٤)$$

مثال (٦,٦) :

أوجد قيمة المتغير العشوائى التابع لتوزيع مربع كاى على ٧٠ درجة والذى تقع على يساره ٠,٩٧٥ من المساحة.

الحل :

$$\text{د} = \text{٧٠}$$

$$\text{أ} = \text{٠,٩٧٥}$$

وباستخدام جدول التوزيع الطبيعى المعيارى يتضح أن :

$$\text{ق}\,(\text{ى}_{\text{٠,٩٧٥}}) = \text{١,٩٦}$$

أما

$$\sqrt{\text{٢ د} - \text{١}} = \sqrt{\text{١٤٠} - \text{١}}$$

$$= \sqrt{\text{١٣٩}}$$

$$\sqrt{\text{٢ د} - \text{١}} = \text{١١,٨}$$

$$\text{ك} = \frac{\text{١}}{\text{٢}}\left(\text{ق (ى)} + \sqrt{\text{٢ د} - \text{١}}\right)^{\text{٢}}$$

$$= \frac{\text{١}}{\text{٢}}\left(\text{١,٩٦} + \text{١١,٨}\right)^{\text{٢}}$$

$$= \frac{\text{١}}{\text{٢}}\left(\text{١٣,٧٦}\right)^{\text{٢}}$$

$$= \frac{\text{١}}{\text{٢}} \times \text{١٨٩,٣}$$

$$= \text{٩٥}$$

٥ - توزيع ف (F Distribution) :

إذا كان المتغير العشوائى $\text{س}_{\text{١}}$ يتبع توزيع مربع كاى على $\text{د}_{\text{١}}$ من درجات الحرية ، وإذا كان المتغير $\text{س}_{\text{٢}}$ يتبع أيضاً توزيعاً لمربع كاى، ولكنه على $\text{د}_{\text{٢}}$ من درجات الحرية فإن :

$$\frac{\text{س}_{\text{١}} / \text{د}_{\text{١}}}{\text{س}_{\text{٢}} / \text{د}_{\text{٢}}} \qquad (\text{١٥})$$

تتبع لتوزيع فيشر (ف) بعدد من درجات حرية يساوى $د_1$ و $د_2$. وتسمى $د_1$ بعدد درجات حرية البسط و $د_2$ عدد درجات حرية المقام ويكتب على النحو التالى :

$$ف_{د_1 ، د_2}.$$

إذاً فتوزيع فيشر (Fisher) هو عبارة عن نسبة توزيعين لمربع كاى، كل مقسوم على درجات حريته، شريطة أن يكونا مستقلين.

مثال (٦,٧) :

أثبت أن نسبة تباين العينة الأولى إلى تباين العينة الثانية المستقلة تتبع توزيع ف، وبين عدد درجات الحرية للتوزيع ، مفترضاً أن تبايني المجتمعين اللذين سحبت منها العينتان متساويان.

الحل :

افرض أن عدد قيم العينة الأولى يساوى $ن_1$ بينما عدد قيم العينة الثانية يساوى $ن_2$ ، وافرض أن $ع^2_1$ و $ع^2_2$ يمثلان تبايني العينة الأولى والثانية على التوالى. كما أن $م^2_1$ و $م^2_2$ هما تباينا المجتمعين اللذين سحبت منها العينتان. باستخدام الخاصية الثالثة لتوزيع مربع كاى يتضح أن :

$$\frac{(ن_1 - ع^2_1)}{م^2_1}$$ تتبع لتوزيع مربع كاى على $(ن_1 - ١)$ درجات حرية

بذلك :

$$\frac{(ن_2 - ع^2_2)}{م^2_2}$$ تتبع لتوزيع مربع كاى على $(ن_2 - ١)$ درجات حرية

وبافتراض أن $م^2_1 = م^2_2$ وبقسمة كل مقدار على عدد درجاته من الحرية يتضح أن :

$$ف_{(ن_1 - ١) ، (ن_2 - ١)} = \frac{\dfrac{[(ن_1 - ١)ع^2_1] / (ن_1 - ١)}{م^2_1}}{\dfrac{[(ن_2 - ١)ع^2_2] / (ن_2 - ١)}{م^2_2}}$$

وبما أن

$$م_{١}^{٢} = م_{٢}^{٢} \qquad \text{فإن :}$$

$$\frac{ع_{١}^{٢}}{ع_{٢}^{٢}} = ف_{(ن_{١} - ١)\text{، }(ن_{٢} - ١)} \qquad (١٦)$$

ولهذه النتيجة استخدامات كثيرة جداً فى اختبارات الفروض، لذلك يسمى توزيع فيشر فى بعض الحالات بتوزيع نسبة التباين (Variance – Ratio Distribution) .

هذا، ولقد دون توزيع (ف) فى جداول (ملحق رقم ٤). وبما أن له معلمين، هما : عدد درجات حرية البسط، وعدد درجات حرية المقام، فلقد دونت درجات البسط على الأعمدة (أفقياً)، ودونت درجات المقام فى أول عمود على اليسار، وعلى مقدمة الجدول نسبة المساحة الواقعة إلى يمين القيمة المعنية. وسوف يتضح استخدام هذه الجداول فى الفصول القادمة.

٦ - توزيع ت (Stdudent t–Distribution) :

إذا كانت ى تمثل متغيراً عشوائياً يتبع التوزيع الطبيعى المعيارى، وإذا كانت س متغيراً عشوائياً مستقلاً عن ى ويتبع توزيع مربع كاى على عدد من درجات الحرية يساوى د فتوزيع (ت) على عدد من درجات الحرية يساوى د أيضاً هو :

$$ت_{د} = \frac{ى}{\sqrt{س / د}} \qquad (١٧)$$

أى أنه نسبة التوزيع الطبيعى المعيارى إلى الجذر التربيعى لتوزيع مربع كاى، مقسوماً على درجات حريته.

مثال (٦،٨) :

أوجد العلاقة بين توزيع ت وتوزيع ف.

الحل :

$$ت_{د} = \frac{ى}{\sqrt{س / د}}$$

$$(ت_{د})^{٢} = \frac{ى^{٢}}{س / د} \qquad (١٨)$$

بما أن مربع التوزيع الطبيعى المعيارى يتبع توزيع مربع كاى على عدد من درجات الحرية يساوى الواحد . .

إذاً :

$$(ت_{د})^{2} = \frac{ى^{2} / ١}{س / د}$$

$$= ف_{١ ، د}$$

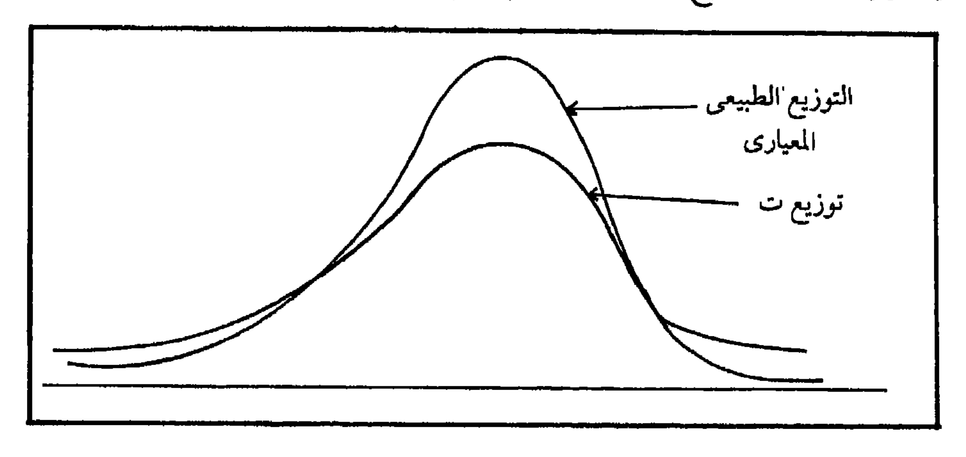
وبمعنى آخر فإن :

$$ت_{د} = \sqrt{ف_{١ ، د}} \qquad (١٩)$$

وهذا يعنى أن $ف_{١ ، د}$، تساوى مربع قيمة $ت_{د}$ ويمكن للقارىء التأكد من ذلك بعد معرفة كيفية استخدام جداول (ت)، التى تعتمد على معلم واحد هو عدد درجات الحرية (ملحق رقم ٢). هذا، وتستخدم جداول ت بنفس طريقة استخدام جداول مربع كاى.

وتجدر الإشارة هنا إلى أن توزيع ت يشبه إلى حد كبير التوزيع الطبيعى المعيارى، إلا أن المساحة حول الصفر أكبر لدى التوزيع الطبيعى ، وبالتالى أصبح أعلى قمة من توزيع ت. كما أن طرفى (ذيلى) التوزيع الطبيعى أقرب إلى المحور الأفقى عند النهاية.

هذا، ويقترب توزيع ت من التوزيع الطبيعي كلما ازداد عدد درجات الحرية، إلى أن يتساوى معه عندما يصبح عدد درجات الحرية كبيراً.



٧ ـ توزيع الوسط الحسابى للعينة (Sampling Distribution of Sample Means) :

إذا سحبت العينات $س_١$ ، $س_٢$ ، $س_٣$ ، . . . . . ، $س_{ن}$ من نفس المجتمع ، وبطريقة تجعل كل عينة مستقلة عن الأخرى (بطريقة عشوائية) ، وإذا كانت الأوساط الحسابية لتلك العينات هى $\bar{س}_١$ ، $\bar{س}_٢$ ، $\bar{س}_٣$ ، . . . . . ، $\bar{س}_{ن}$ ، فالوسط الحسابى لكل عينة يمثل تقديراً لوسط المجتمع (و). وسواء كان التوزيع الخاص بالمجتمع طبيعياً أو غير طبيعى ، فإن الوسط الحسابى لأوساط العينات ($\bar{ن}$) التى يساوى حجم كل منها ن هو :

$$\bar{س} = \frac{\bar{س}_١ + \bar{س}_٢ + \bar{س}_٣ + \ldots + \bar{س}_{ن}}{\bar{ن}}$$

يتبع توزيعاً وسطه (و) وتباينه $\frac{م^٢}{ن}$ حيث $م^٢$ هى تباين المجتمع الذى سحبت منه العينات ويسمى الانحراف المعيارى لتوزيع الوسط الحسابى للعينة ($\frac{م}{\sqrt{ن}}$) بالخطأ المعيارى للوسط (Standard Error of the Mean) ويلاحظ هنا أن $م_{\bar{س}}$ (الخطأ المعيارى للوسط) أقل من الانحراف المعيارى لأى وسط من أوساط العينات . كما أن الخطأ المعيارى يتناقص بزيادة ن ، أى بزيادة تكرار المعاينة . وهذا يعنى أن زيادة حجم العينات المسحوبة يزيد من دقة تقدير وسطها الحسابى ($\bar{س}$) لوسط المجتمع (و).

مثال (٦,٩) :

جلس ٦ طلاب للاختبار فى مادة الرياضيات وكانت درجاتهم :

٦٨ ؛ ٧٠ ؛ ٧٢ ؛ ٧٤ ؛ ٧٦ ؛ ٧٨

أوجد جميع العينات الثنائية الممكنة ، و أوساطها ، والوسط الحسابى لها .

الحل :

يوضح الجدول التالى كل العينات الثنائية التى يمكن سحبها والتى تساوى ٣٦ عينة .

جدول رقم (١) : العينات وأوساطها الحسابية

العينة		الوسط الحسابى (س̄ ر)	العينة		الوسط الحسابى (س̄ ر)
٦٨	٦٨	٦٨	٦٨	٧٤	٧١
٧٠	٦٨	٦٩	٧٠	٧٤	٧٢
٧٢	٦٨	٧٠	٧٢	٧٤	٧٣
٧٤	٦٨	٧١	٧٤	٧٤	٧٤
٧٦	٦٨	٧٢	٧٦	٧٤	٧٥
٧٨	٦٨	٧٣	٧٨	٧٤	٧٦
٦٨	٧٠	٦٩	٦٨	٧٦	٧٢
٧٠	٧٠	٧٠	٧٠	٧٦	٧٣
٧٢	٧٠	٧١	٧٢	٧٦	٧٤
٧٤	٧٠	٧٢	٧٤	٧٦	٧٥
٧٦	٧٠	٧٣	٧٦	٧٦	٧٦
٧٨	٧٠	٧٤	٧٨	٧٦	٧٧
٦٨	٧٢	٧٠	٦٨	٧٨	٧٣
٧٠	٧٢	٧١	٧٠	٧٨	٧٤
٧٢	٧٢	٧٢	٧٢	٧٨	٧٥
٧٤	٧٢	٧٣	٧٤	٧٨	٧٦
٧٦	٧٢	٧٤	٧٦	٧٨	٧٧
٧٨	٧٢	٧٥	٧٨	٧٨	٧٨

فالوسط الحسابى للعينات :

س̄ = (٦٨ + ٦٩ + ٧٠ + ... + ٧٦ + ٧٧ + ٧٨) / ٣٦

= ٢٦٢٨ / ٣٦

= ٧٣

أما الخطأ المعيارى للوسط فى هذه الحالة فهو :

$$م_{\bar{س}} = \sqrt{\frac{\sum \bar{س}_ر^2 - \frac{(\sum \bar{س}_ر)^2}{ن}}{ن}} \quad (٢٠)$$

$$= ٢,٤$$

أما الوسط الحسابى للمجتمع فهو :

$$و = \frac{٦٨ + ٧٠ + ٧٢ + ٧٤ + ٧٦ + ٧٨}{٦}$$

$$= \frac{٤٣٨}{٦}$$

$$= ٧٣$$

بينما الانحراف المعيارى للمجتمع هو :

$$م = \left(\frac{\sum س_ر^2 - \frac{(\sum س_ر)^2}{ن}}{ن}\right)^{\frac{1}{2}} \quad (٢١)$$

$$= ٣,٤$$

وبما أن حجم العينة هنا = ٢ فإن :

$$\frac{م}{\sqrt{ن}} = \frac{٣,٤}{\sqrt{٢}}$$

$$= ٢,٤$$

البرنامج التالى يقوم بحساب الآتى لمجموعة من البيانات :

- جميع العينات الممكنة .
- الوسط الحسابى لكل عينة .
- الوسط الحسابى للعينات كلها .
- الخطأ المعيارى لوسط العينة .
- الوسط الحسابى للمجتمع .
- الانحراف المعيارى لوسط المجتمع .
- الخطأ المعيارى لوسط المجتمع .

بالمعادلات التالية :

$$X_1 = S/N_1 = \text{الوسط الحسابى للمجتمع}$$

$$V_1 = \sqrt{\left(R - \frac{(S)^2}{N_1} \right) / N_1} = \text{الانحراف المعيارى للمجتمع}$$

$$X_2 = S/N_2 = \text{الوسط الحسابى لأوساط العينات}$$

$$V_2 = \sqrt{\left(R - \frac{S^2}{N_2} \right) / N_2} \quad \text{الانحراف للأوساط}$$

$$E = V_1 / \sqrt{N_3} \quad \text{الخطأ المعيارى للوسط}$$

```
10  REM  برنامج لايجاد جميع العينات الممكنه واوساطها الحسابيد
20  DIM A(6),T(36)
30  READ N1,N2,N3
35  REM        حجم العينه=N3، عدد العينات=N2، عدد المشاهدات=N1
40  MAT READ A
50  PRINT USING 70
60  PRINT USING 80
70  :                          البيانات الاصليه
80  :                          ________________
90  :                                ###
100 :                          الوسط الحسابى للمجتمع = ###
110 :                          الانحراف المغيارى للمجتمع = ###.##
115 PRINT
120 S=0
130 R=0
140  FOR I=1 TO N1
150  PRINT USING 90,A(I)
160  S=S+A(I)
170  R=R+A(I)**2
180 NEXT I
190 X1=S/N1
200 V1=SQR((R-S**2/N1)/N1)
205 PRINT
210 PRINT USING 100,X1
220 PRINT USING 110,V1
230 S=0
240 R=0
250 K=1
260:                    العينـــات واوساطها الحسابيد
270:                    ______________________________
280:                    الوسط الحسابى          العينـــه
290:                    _____________          ________
300:                         ###               ###  ###
310:                    الوسط الحسابى لاوساط العينات = ###
320:                    الانحراف المغيارى لاوساط العينات = ###.##
325:                    الخطا المعيارى للوسط = ###.##
328 PRINT
330 PRINT USING 260
340 PRINT USING 270
350 PRINT USING 280
360 PRINT USING 290
370 FOR I=1 TO N1
380  FOR J=1 TO N1
```

```
390    T(K)=(A(I)+A(J))/2
400    PRINT USING 300,T(K),A(J),A(I)
410    S=S+T(K)
420    R=R+T(K)**2
430    K=K+1
440   NEXT J
450 NEXT I
460 X2=S/N2
470 V2=SQR((R-S**2/N2)/N2)
480 E=V1/SQR(N3)
485 PRINT
490 PRINT USING 310,X2
500 PRINT USING 320,V2
510 PRINT USING 325,E
520 DATA 6,36,2,68,70,72,74,76,78
530 END
```

المحرجات

البيانات الاصلية

68
70
72
74
76
78

الوسط الحسابى للمجتمع = 73
الانحراف المعيارى للمجتمع = 3.416

العينات واوساطها الحسابية

العينة		الوسط الحسابى
68	68	68
68	70	69
68	72	70
68	74	71
68	76	72
68	78	73
70	68	69
70	70	70
70	72	71
70	74	72
70	76	73
70	78	74
72	68	70
72	70	71
72	72	72
72	74	73
72	76	74
72	78	75
74	68	71
74	70	72
74	72	73
74	74	74
74	76	75
74	78	76
76	68	72
76	70	73
76	72	74
76	74	75
76	76	76
76	78	77
78	68	73
78	70	74
78	72	75
78	74	76
78	76	77
78	78	78

الوسط الحسابى لاوساط العينات = 73
الانحراف المعيارى لاوساط العينات = 2.415
الخطأ المعيارى للوسط = 2.415

وهكذا أصبح وسط العينة مطابقاً تماماً لوسط المجتمع (٧٣)، بينما تساوى الخطأ المعيارى مع الانحراف المعيارى للمجتمع . هذا، وتجدر الإشارة إلى أن تباين وسط العينة يكون أقرب للقيمة

$$\frac{نً - ن}{نً - ١} \quad \frac{م^{٢}}{ن}$$

حيث نً هى حجم المجتمع . ويسمى معامل تباين الوسط بمعامل التصحيح الذى لا يستخدم إلا إذا كان حجم المجتمع صغيراً جداً . أما إذا كان حجم المجتمع كبيرا فإن قيمته تقترب من الواحد .

وأما التوزيع التكرارى للأوساط الحسابية للعينات السالفة الذكر فهو كما يلى :

جدول (٢)

التوزيع التكرارى للأوساط الحسابية للعينات الثنائية البالغ عددها ٣٦ عينة .

الأوساط الحسابية (سَ ر)	التكـرار (ك ر)
٦٨	١
٦٩	٢
٧٠	٣
٧١	٤
٧٢	٥
٧٣	٦
٧٤	٥
٧٥	٤
٧٦	٣
٧٧	٢
٧٨	١
المجموع	٣٦

التوزيع التكرارى السالف الذكر يمثل توزيعاً متماثلاً حول الوسط، بمعامل التواء يساوى الصفر، ومعامل تفرطح يساوى ٢,٤، فهو بذلك لا يختلف اختلافاً جوهرياً عن التوزيع الطبيعى الأمثل. وليس ذلك مجرد صدفة، أو حالة خاصة، ولكن توزيع أوساط العينات يزداد قرباً من التوزيع الطبيعى، كلما ازداد عدد العينات، أى كلما ازداد عدد الأوساط (ن). وعليه تكون :

$$\bar{س} \sim ط\ (و ،\ \frac{م^{٢}}{ن}) \qquad (٢٢)$$

حيث (~) تعنى أن المتغير يتبع التوزيع، و (ط) تعنى «طبيعى». ويؤدى ذلك إلى اعتبار أن :

$$\frac{\bar{س} - و}{م / \sqrt{ن}} \sim ط\ (٠ ، ١) \qquad (٢٣)$$

إلا أن اقتراب توزيع المتغيرات العشوائية من التوزيع الطبيعى المعيارى ليس مقصوراً على وسط أوساط العينات، فلقد ثبت أن هذه الخاصية يمكن تعميمها على جميع الأوساط الحسابية إذا كان عدد المتغيرات (ن) كبيراً. لذلك فقد جاءت صياغة النهاية المركزية (Central Limit Theorem)، أو النظرية الأساسية لتقارب العينات على النحو الآتى :

نظرية (١) :

«إذا كانت $س_{١}$، $س_{٢}$، $س_{٣}$، . . . .، $س_{ن}$ متغيرات عشوائية تتبع توزيعاً واحداً، وسطه (و)، وانحرافه المعيارى (م)، وإذا كانت ($\bar{س}$) هى الوسط الحسابى لتلك المتغيرات العشوائية، فالتوزيع النهائى ($ن \leftarrow \infty$) للمتغير :

$$\frac{\bar{س} - و}{م / \sqrt{ن}}$$

هو التوزيع الطبيعى المعيارى».

وتجدر الإشارة هنا إلى أن مقدار (ن) الذى يصبح بموجبه التوزيع طبيعياً ليس محدداً؛ إذ أنه يزيد كلما كان التوزيع الأصلى للمتغيرات بعيداً عن التوزيع الطبيعى، إلا أن أكثر المراجع تعتبر ٣٠ حداً أدنى لقيمة (ن).

بيد أن هناك حالات كثيرة يكون فيها تباين المجتمع (م^{2}) غير معلوم، وعندها يلجأ الباحث لاستخدام العلاقة بين التوزيع الطبيعي وتوزيع مربع كاى؛ ذلك لأن توزيع تباين العينة(ع^{2}) الذى يعتبر مقدراً لتباين المجتمع يتبع توزيع مربع كاى.

وبما أن :

$$\frac{\text{ى}}{\sqrt{\text{ك}^{2} / \text{ن}}} \sim \text{ت}_{(\text{ن})} \qquad (٢٤)$$

حيث ى ~ ط (. ، ١)

ن = عدد درجات حرية توزيع مربع كاى (ك^{2}) ومن ثم عدد درجات توزيع (ت).

وبما أن :

$$\frac{(\text{ن} - ١)\,\text{ع}^{2}}{\text{م}^{2}} \sim \text{ك}^{2}_{(\text{ن} - ١)} \qquad (٢٥)$$

فإن :

$$\frac{\left(\dfrac{\bar{\text{س}} - \text{و}}{\text{م} / \sqrt{\text{ن}}}\right)}{\left(\dfrac{(\text{ن} - ١)\,\text{ع}^{2}}{\text{م}^{2}\,(\text{ن} - ١)}\right)^{\frac{1}{2}}} \sim \text{ت}_{(\text{ن} - ١)} \qquad (٢٦)$$

إذاً :

$$\frac{\bar{\text{س}} - \text{و}}{\text{ع} / \sqrt{\text{ن}}} \sim \text{ت}_{(\text{ن} - ١)}$$

وهذا يعنى أنه إذا استبدل تباين العينة عوضاً عن تباين المجتمع، فسوف يصبح التوزيع تابعاً لتوزيع ت على (ن ـ ١) درجات حرية.

٨ ـ توزيع مجموع الوسطين أو الفرق بينهما :

نظرية (٢) :

«إذا تم سحب عينتين مستقلتين من مجتمعين موزعين توزيعاً طبيعياً، الأول : بوسط (و_{1}) وتباين (م^{2}_{1})، والثانى : بوسط (و_{2}) وتباين (م^{2}_{2}). وإذا كان حجم العينة الأولى (ن_{1})

ووسطها $(\bar{س})^{٢}$، بينما كان حجم العينة الثانية $(ن_{٢})$ ووسطها الحسابى $(\bar{ص})$.

فإن :

$$\frac{\bar{س} + \bar{ص} - (و_{١} + و_{٢})}{\sqrt{\frac{م_{١}}{ن_{١}} + \frac{م_{٢}}{ن_{٢}}}} \sim ط(٠ ، ١)$$

أى أن مجموع الوسطين موزع توزيعاً طبيعياً بوسط يساوى مجموع وسطين، وتباين يساوى مجموع تباينى الوسطين؛ ذلك لأن :

$$\bar{س} \sim ط\left(و_{١} ، \frac{م^{٢}_{١}}{ن_{١}}\right)$$

$$\bar{ص} \sim ط\left(و_{٢} ، \frac{م^{٢}_{٢}}{ن_{٢}}\right)$$

$$\bar{س} + \bar{ص} \sim ط\left(و_{١} + و_{٢} ، \frac{م^{٢}_{١}}{ن_{١}} + \frac{م^{٢}_{٢}}{ن_{٢}}\right)$$

أما فى حالة الفرق بين الوسطين فإن :

$$\bar{س} - \bar{ص} \sim ط\left(و_{١} - و_{٢} ، \frac{م^{٢}_{١}}{ن_{١}} + \frac{م^{٢}_{٢}}{ن_{٢}}\right)$$

وعليه :

$$\frac{(\bar{س} - \bar{ص}) - (و_{١} - و_{٢})}{\left(\frac{م^{٢}_{١}}{ن_{١}} + \frac{م^{٢}_{٢}}{ن_{٢}}\right)^{\frac{١}{٢}}} \sim ط(٠ ، ١) \qquad (٢٧)$$

هذا، ويمكن التعويض عن $م_{١}$ و$م_{٢}$ بتباينى العينتين $ع_{١}$ و$ع_{٢}$ على التوالى، إذا كان حجم العينة كبيراً فى الحالتين، بينما كان تباينا العينتين مجهولين. وبذلك يصبح توزيع مجموع الوسطين على النحو التالى :

$$\frac{(\bar{س} + \bar{ص}) - (و_{١} + و_{٢})}{\sqrt{\frac{ع^{٢}_{١}}{ن_{١}} + \frac{ع^{٢}_{٢}}{ن_{٢}}}} \sim ط(٠ ، ١) \qquad (٢٨)$$

أما إذا كانت $م_١ = م_٢ = م$ ، أى أن تبايني المجتمعين متساويان ، فتوزيع مجموعهما هو :

$$\frac{(\bar{س} + \bar{ص}) - (و_١ + و_٢)}{م \sqrt{\frac{١}{ن_١} + \frac{١}{ن_٢}}} \qquad (٢٩)$$

بيد أن هناك حالات كثيرة يكون فيها حجما العينتين صغيرين بدرجة لا يمكن معها التعويض بتبا ينيهما عن تبايني المجتمعين. وبافتراض أن تبايني المجتمعين متساويان ($م_١^٢ = م_٢^٢ = م^٢$)، واعتماداً على توزيع مربع كاى وخواصه التى تحدد أن :

$$\frac{(ن_١ - ١) ع_١^٢}{م^٢} \sim ك^٢_{(ن_١ - ١)} \qquad (٣٠)$$

$$\frac{(ن_٢ - ١) ع_٢^٢}{م^٢} \sim ك^٢_{(ن_٢ - ١)} \qquad (٣١)$$

وبما أن :

$$\frac{(ن_١ - ١) ع_١^٢}{م^٢} + \frac{(ن_٢ - ١) ع_٢^٢}{م^٢} \sim ك^٢_{(ن_١ + ن_٢ - ٢)} \qquad (٣٢)$$

واعتماداً على أن الوسط الحسابى والتباين لنفس العينة يكونان مستقلين بعضهما عن بعض فإن :

$$\frac{\left(\frac{\bar{س} + \bar{ص} - (و_١ + و_٢)}{م \sqrt{\frac{١}{ن_١} + \frac{١}{ن_٢}}}\right)}{\left(\frac{\frac{(ن_١ - ١) ع_١^٢ + (ن_٢ - ١) ع_٢^٢}{م^٢}}{(ن_١ + ن_٢ - ٢)}\right)^{\frac{١}{٢}}} = \frac{ط(٠ ، ١)}{\sqrt{ك^٢ / (ن_١ + ن_٢ - ٢)}} \qquad (٣٣)$$

وتكون فرضية العدم هنا إما على النحو الآتى :

$ف_1 : و_1 > و$

أو :

$ف_1 : و_1 < و_2$

أو بمعنى آخر :

$ف_1 - ف_2 > صفر$

أو :

$ف_2 - ف_1 < صفر$

وفى جميع هذه الحالات تكون على طرف واحد، وتعنى أن هناك وسطاً أكبر من الآخر.

كذلك قد تكون الفرضية البديلة على النحو التالى :

$ف_1 : و_1 \neq و_2$

وهى فرضية ذات طرفين تعنى أن هناك فرقاً جوهرياً بين الوسطين.

أما إذا كان الفرق المحدد بفرضية العدم غير الصفر - يساوى ل مثلاً حيث ل قيمة معلومة - بمعنى أن :

$ف_0 : و_1 - و_2 = ل$ وهى فرضية عدم بسيطة.

أو كانت مزدوجة على النحو الآتى :

$ف_0 : و_1 - و_2 > ل$

فقد تكون الفرضية البديلة على النحو الاتى :

$ف_1 : و_1 - و_2 > ل$

أو

$ف_1 : و_1 - و_2 < ل$

أو

$ف_1 : و_1 - و_2 \neq ل$

أى أن :

$$\frac{\text{س} - \text{ن ح}}{\sqrt{\text{ن ح}\,(١ - \text{ح})}} \sim \text{ط}\,(٠ ، ١) \qquad (٣٨)$$

أما إذا كان الهدف هو تحديد نسبة الذين يتمتعون بتلك الصفة بدلاً من عددهم للاستدلال بنسبة العينة على نسبة المجتمع ، فإذا كانت $\bar{\text{ح}}$ تعنى نسبة العينة فإن :

$$\bar{\text{ح}} = \frac{\text{س}}{\text{ن}}$$

وبقسمة بسط ومقام المعادلة على ن يصبح التوزيع :

$$\frac{\bar{\text{ح}} - \text{ح}}{\sqrt{\frac{\text{ح}\,(١ - \text{ح})}{\text{ن}}}} \sim \text{ط}\,(٠ ، ١) \qquad (٣٩)$$

وهى شبيهة بتوزيع وسط العينة بعد استبدال $\bar{\text{ح}}$ مكان $\bar{\text{س}}$ و ح (١ - ح) مكان $\text{م}^{٢}$. لذلك يمكن استخدام معادلتى المجموع، والفرق بين وسطين، لتقدير توزيع المجموع والفرق بين نسبتين، وذلك على النحو التالى :

$$\frac{(\bar{\text{ح}}_{١} + \bar{\text{ح}}_{٢}) - (\text{ح}_{١} + \text{ح}_{٢})}{\sqrt{\frac{\text{م}_{١}^{٢}}{\text{ن}_{١}} + \frac{\text{م}_{٢}^{٢}}{\text{ن}_{٢}}}} \sim \text{ط}\,(٠ ، ١) \qquad (٤٠)$$

حيث

$$\text{م}_{١}^{٢} = \text{ح}_{١}\,(١ - \text{ح}_{١})$$

$$\text{م}_{٢}^{٢} = \text{ح}_{٢}\,(١ - \text{ح}_{٢})$$

أما توزيع الفرق بين نسبتين فهو :

$$\frac{(\bar{\text{ح}}_{١} - \text{ح}_{٢}^{٢}) - (\text{ح}_{١} - \text{ح}_{٢})}{\sqrt{\frac{\text{م}_{١}^{٢}}{\text{ن}_{١}} + \frac{\text{م}_{٢}^{٢}}{\text{ن}_{٢}}}} \sim \text{ط}\,(٠ ، ١) \qquad (٤١)$$

وبما أن مربع التوزيع الطبيعى المعيارى هو مربع كاى على درجة حرية واحدة.

فإن :

$$\frac{(س - ن ح)^{2}}{ن ح (1 - ح)} \sim ك^{2}_{1} \qquad (42)$$

نظرية (٤):

إذا كانت $س_{1}$، $س_{2}$، $س_{3}$، . . . . . . ، $س_{ل}$، هى متغيرات عشوائية متعددة على التوزيع ذى الحدين، وإذا كانت ن، $ح_{1}$، $ح_{2}$، $ح_{3}$، . . . . . . ، $ح_{ل}$، هى معالم التوزيعات

حيث

$$ح_{1} + ح_{2} + ح_{3} + \cdots + ح_{ل} = 1 \qquad (43)$$

فإن :

$$\sum_{ر=1}^{ل} \frac{(س_{ر} - ن ح_{ر})^{2}}{ن ح_{ر}} \sim ك^{2}_{(ل - 1)} \qquad (44)$$

فإذا كانت

$$ل = 2$$

بينما

$س_{1}$ = عدد حالات النجاح

و

$س_{2}$ = عدد حالات الفشل

فإن

$$ح_{2} = 1 - ح_{1}$$

كما أن :

$$س_{2} = ن - ح_{1}$$

وعليه :

$$\sum_{\text{ر}=١}^{٢} \frac{(\text{س}_{\text{ر}} - \text{ن ح}_{\text{ر}})^{٢}}{\text{ن ح}_{\text{ر}}} = \frac{(\text{س}_{١} - \text{ن ح}_{١})^{٢}}{\text{ن ح}_{١}} + \frac{(\text{س}_{٢} - \text{ن ح}_{٢})^{٢}}{\text{ن ح}_{٢}}$$

$$= \frac{(\text{س}_{١} - \text{ن ح}_{١})^{٢}}{\text{ن ح}_{١}\ (١ - \text{ح}_{١})} \sim \text{ك}^{٢}_{١} \qquad (٤٦)$$

وبما أن س ر هى دائماً عدد حالات النجاح التى تحققت، أو عدد تكرار حالات النجاح (ك ر) بينما ن ح ر تمثل عدد حالات النجاح المتوقعة فى المجتمع (كَ ر)، لذلك فإن :

$$\frac{\sum_{\text{ر}=١}^{\text{ن}} (\text{ك}_{\text{ر}} - \text{كَ}_{\text{ر}})^{٢}}{\text{كَ}_{\text{ر}}} \sim \text{ك}^{٢}_{(\text{ن}-١)} \qquad (٤٧)$$

تمارين

(١) إذا كانت ح تعنى احتمالاً، بينما تعنى ى القيمة المعيارية، فأوجد ما يلى باستخدام جدول التوزيع الطبيعى :

(أ) ح (ى > ١).
(ب) ح (ى < ١).
(جـ) ح (١ > ى > ٢).
(د) ح (|ى| > ½).

(٢) استخدم نفس رموز السؤال الأول وجدول التوزيع الطبيعى لإيجاد قيمة و إذا كان :

(أ) ح (ى < و) = ٠,٠٢٥
(ب) أثبت أن ح (ى < و) = ح (ى < ـ و)، لأى قيمة بين الصفر والثلاثة.
(ج) ح (|ى| > و) = ٠,٩٥

(٣) يفترض أن توزيع رواتب العاملين بإحدى المؤسسات يتبع التوزيع الطبيعى بوسط حسابى ٦٠٠٠ ريال، وانحراف معيارى ٣٠٠٠ ريال. فإذا تم اختيار أحد العاملين بتلك المؤسسة اختياراً عشوائياً فما احتمال أن يكون راتبه :

(أ) ٦٠٠٠ ريال.
(ب) ٩٠٠٠ ريال.
(جـ) ٣٠٠٠ ريال.
(د) أكثر من ٥٠٠٠ ريال.
(هـ) أقل من ٥٠٠٠ ريال.
(و) أكثر من ٩٠٠٠ ريال.
(ز) أقل من ٩٠٠٠ ريال.
(ح) بين ٤٠٠٠ ريال و ١٠٠٠٠ ريال.

(٤) ينتج أحد المصانع قضباناً حديدية حمولة القطعة منها ١٥٠ كيلوجراماً، بانحراف معيارى ١٠ كيلوجرامات. فما هو عدد القطع من بين ١٠٠٠ قطعة المتوقع أن تكون حمولته بين ١٤٥ و ١٦٠ كيلوجراماً؟

(٥) نسبة المراجعين (لأحد المستوصفات) المصابين بأحد أمراض الباطنية تساوى ١٠٪، فإذا اختيرت عينة عشوائية قوامها ١٠ مراجعين فما احتمال أن تكون :

(أ) كلها خالية من ذلك المرض.
(ب) كلها مصابة بذلك المرض.
(جـ) عدد المصابين شخصين فقط.
(د) عدد المصابين ٣ أشخاص فقط.
(هـ) عدد المصابين بين ٤ و ٧ أشخاص.
(و) ما هو عدد المصابين المتوقع من بين أفراد العينة؟

(٦) اذكر العلاقة بين كل توزيعين :

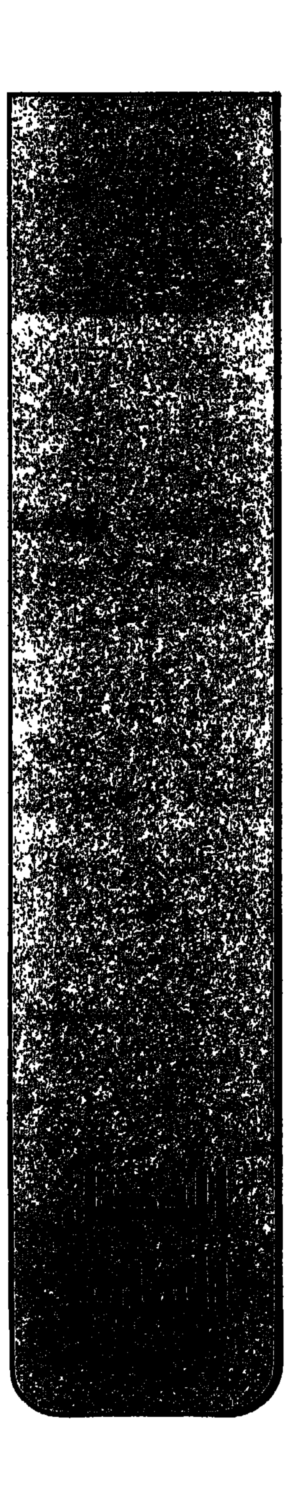
(أ) ذى الحدين والطبيعى.
(ب) الطبيعى ومربع كاى.
(جـ) مربع كاى و ت.
(د) ت و ف.
(هـ) ف ومربع كاى.

(٧) إذا كان و متغيراً عشوائياً يتبع توزيع مربع كاى على ١٥ درجة حرية، فأوجد :

(أ) ح (س ≥ و) = ٠,٩٠
(ب) ح (س ≥ و) = ٠,٠٠٥
(جـ) ح (س ≥ و) = ٠,٩٩

(٨) كانت أوزان المدافعين الأربعة لإحدى الفرق كما يلى :
٧٠ كيلوجراماً ، ٧٤ كيلوجراماً ، ٦٨ كيلوجراماً ، ٦٤ كيلوجراماً.
أوجد جميع العينات الثنائية الممكنة، وأوساطها الحسابية، والوسط الحسابى لها جميعاً.

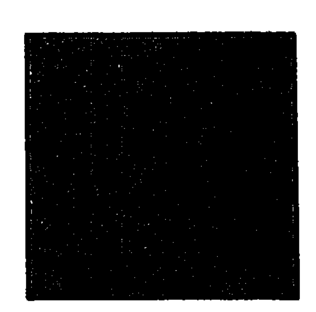
(٩) ما هى أهم مزايا توزيع ت ومتى يستخدم؟
(١٠) متى يستخدم توزيع ف؟



فترات الثقة

CONFIDENCE INTERVALS







فترات الثقة

(Confidence Intervals)

١ ـ الاستدلال الإحصائى (Statistical Inference) :

الاستدلال الإحصائى هو نوع من أنواع اتخاذ القرارات فى ظروف عدم التأكد، ويهدف الاستدلال الإحصائى إلى الوصول للقرار الأسلم، اعتماداً على الاحتمالات، أى بزيادة احتمال أن يكون القرار سليماً. وبما أنه لا يمكن التأكد بصورة قاطعة من سلامة أو عدم سلامة القرار المتخذ فى ظروف عدم التأكد، فقد أصبح الاستدلال الإحصائى هو أفضل البدائل المتاحة أمام الإنسان فى المجالات التى يمكن استخدامه فيها. هذا، وتنقسم مجالات استخدام الاستدلال الإحصائى إلى ثلاثة أقسام، وهى :

١ ـ تقدير النقطة.
٢ ـ تقدير الفترة.
٣ ـ اختبارات الفرضيات.

والتقدير بنوعيه السابقين هو اختيار قيمة محددة كبديل لمعلمة مجهولة، وهذا البديل هو واحد من سلسلة متصلة للبدائل الممكنة باحتمالات مختلفة، أما اختبار الفرضية فيعنى فى مجمله قبول أو رفض قيمة محددة ـ أو عدة قيم ـ خاصة بمعلمة معينة. هذا، وتختلف القيود المستخدمة فى التقدير عن تلك التى تستخدم لاختبار الفرضية.

اهتمت الفصول السابقة بتقدير معالم المجتمع بدلالات القيم العينية، فقدر وسط المجتمع (و) بالوسط الحسابى للعينة (سَ)، وقدر تباين المجتمع (م ٢) بتباين العينة (ع ٢). هذا، ولقد تمثلت دلالات القيم العينية بالعزوم، كذلك قدر وسط المجتمع وتباينه فى حالة معرفة التوزيع الإحصائى الذى تتبعه المتغيرات العشوائية. ويسمى ذلك النوع من التقديرات

بالتقـدير بنقطة (Point Estimation) ؛ لأنه يعتمد على اختيار نقطة واحدة من سلسلة من النقاط لتقدير أحد معالم المجتمع. أما تقدير الفترة - كما يدل على ذلك اسمه - فهو يتعلق بتحديد فترة تسمى فترة الثقة يرجح ، وباحتمال كبير، أن يكون المعلم محصوراً بين تلك الحدود.

وتنقسم فترات أو حدود الثقة إلى ثلاثة أقسام رئيسية، وهى :

١ - فترات الثقة للأوساط.
٢ - فترات الثقة للنسب.
٣ - فترات الثقة للتباينات.

هذا، وسوف يتم استعراض كل قسم من هذه الأقسام فى هذا الفصل لينفرد الفصل التالى باختبار الفرضيات.

٢ - فترات الثقة للأوساط (Confidence Intervals for Means) :

بالرغم من أن وسط العينة (سَ) يمثل تقديراً جيداً لوسط المجتمع (و)، إلا أنه لا يساويه تماماً، فقد اتضح من الفصل الماضى أن العينات المستقلة المسحوبة من نفس المجتمع تنتهى فى أكثر الحالات إلى تقديرات مختلفة لنفس المعلمة (Parameter) .

ولقد اتضح من الفصل الماضى أن الوسط الحسابى للعينات يتبع التوزيع :

$$\text{سَ} \sim \text{ط}\left(\text{و} ، \frac{\text{م}^2}{\text{ن}}\right) \qquad (١)$$

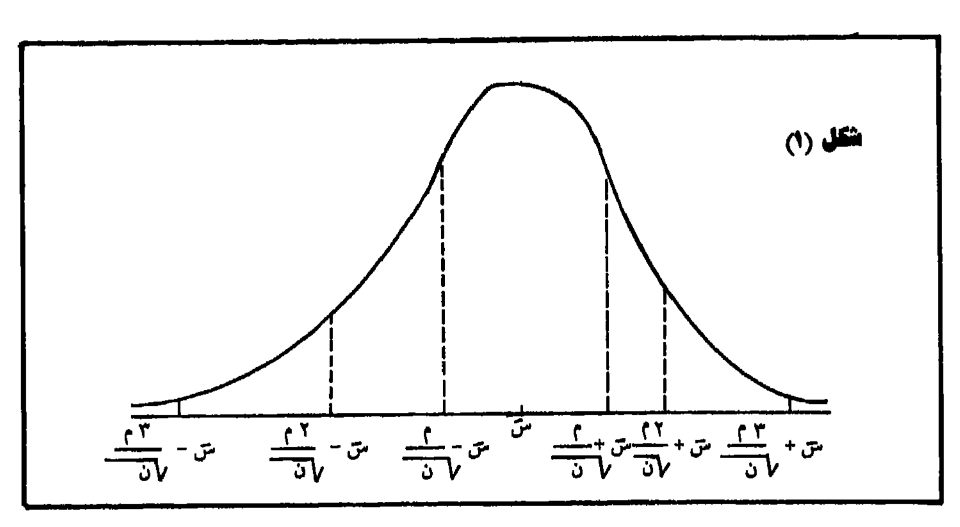
حيث ن هو حجم العينة.

وهذا يعنى أن :

$$\frac{\text{سَ} - \text{و}}{\text{م} / \sqrt{\text{ن}}} \sim \text{س ط} (٠ ، ١) \qquad (٢)$$

وعليه فمن الجائز القول بأن (و) ربما تنحصر فى الفترة سَ ± م/$\sqrt{\text{ن}}$ كما أن هناك احتمالاً أكبر بأن تكون الفترة هى سَ ± ٢ م/$\sqrt{\text{ن}}$، وهكذا.

إلا أن هذه الحدود تختلف باختلاف أوساط العينات فى حالة تساوى الحجم (ن)، بيد أن التوزيع الخاص بوسط العينات هو توزيع طبيعى حسب ما سلف ذكره. إذاً فهناك ٦٨,٣٪ من الأوساط تنحصر فى الفترة سَ ± م/$\sqrt{\text{ن}}$ كما أن هناك ٩٥٪ من الأوساط تنحصر فى الفترة

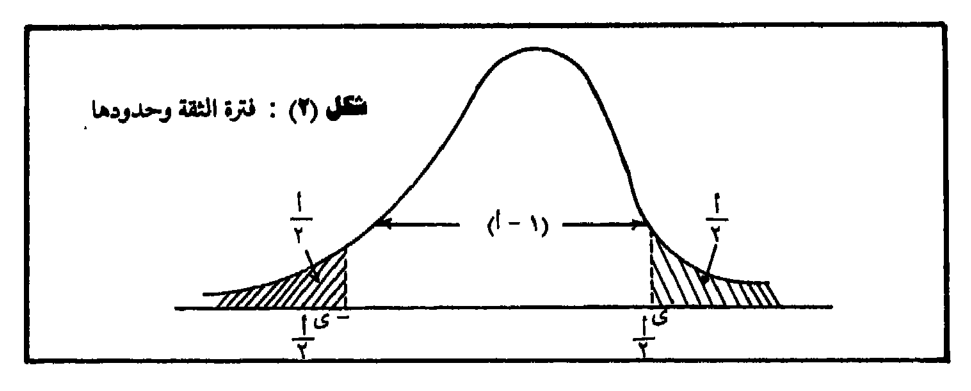


سَ ± ١,٩٦ م/√ن وهناك ٩٩٪ من الأوساط تنحصر فى الفترة سَ ± ٢,٥٧ م/√ن. انظر الشكل (١) السابق. إلا أنه لا يمكن تحديد الفترة التى تنحصر فيها جميع الأوساط وبالتالى وسط المجتمع ؛ ذلك لأن طرفى التوزيع الطبيعى يمتدان إلى ما لا نهاية، كما أنه ليس أمراً منطقياً أن يتم سحب جميع العينات الممكنة عند إجراء كل دراسة.

إذا كانت

$$\text{ى} = \frac{\text{سَ} - \text{و}}{\text{م}/\sqrt{\text{ن}}} \qquad (٣)$$

فإن ى قيمة معيارية. وأما إذا كانت ى ١ - أ/٢ تعنى أن المساحة المحصورة بين ى أ/٢ وما لا نهاية تساوى أ/٢ فإن ى (١ - أ/٢) هى عدد الانحرافات التى تنحصر بينها ١٠٠ (١ - أ/٢)٪ من القيم. إذاً فهناك ١٠٠ (١ - أ)٪ من المساحة تنحصر بين ى أ/٢ و-ى أ/٢ بسبب تشابه التوزيع الطبيعى (انظر الشكل (٢) التالى).



هذا، وتسمى الفترة ١٠٠(١ ـ أ)٪ بفترة الثقة (Confidence Interval) وتسمى نقطتا نهايتها بحدود الثقة (Confidence Limits). وبذلك تكون فترة الثقة بمستوى ١٠٠(١ ـ أ)٪ لوسط المجتمع هى النسبة المئوية من الأوساط الحسابية التى تنحصر بين حدى تلك الفترة عند تكرار تجربة سحب عينة بحجم ن عدة مرات. وبناء عليه تكون :

$$-ى_{\frac{أ}{٢}} < \frac{\bar{س} - و}{م/\sqrt{ن}} < +ى_{\frac{أ}{٢}} \qquad (٤)$$

وهذا يعنى أن :

$$-ى_{\frac{أ}{٢}}\frac{م}{\sqrt{ن}} - \bar{س} < -و < ى_{\frac{أ}{٢}}\frac{م}{\sqrt{ن}} - \bar{س}$$

وبالضرب فى (ـ ١) تكون النتيجة النهائية على النحو التالى :

$$\bar{س}\ ى_{\frac{أ}{٢}}\frac{م}{\sqrt{ن}} < و < \bar{س} + ى_{\frac{أ}{٢}}\frac{م}{\sqrt{ن}} \qquad (٥)$$

وهى أيضاً :

$$\bar{س} - ى_{(١-\frac{أ}{٢})}\frac{م}{\sqrt{ن}} < و < \bar{س} + ى_{(١-\frac{أ}{٢})}\frac{م}{\sqrt{ن}} \qquad (٦)$$

فإذا كانت درجة الثقة هى ٩٠٪ أو ٩٥٪ أو ٩٩٪ فإن $\frac{أ}{٢}$ تساوى ٠,٠٥ أو ٠,٠٢٥ أو ٠,٠٠٥ على التوالى. أما $ى_{(١-\frac{أ}{٢})}$ فتكون على النحو التالى (راجع جدول التوزيع الطبيعى بالملحق) :

$$ى_{٠,٩٠} = ١,٦٤$$

$$ى_{٠,٩٥} = ١,٩٦$$

$$ى_{٠,٩٩} = ٢,٥٧$$

مثال (٧,١) :

أخذت عينة من درجات ٢٥ طالباً فى مادة الرياضيات، فوجد أن وسطها الحسابى ٧٢ درجة، فما فترة ٩٠٪ ثقة للوسط لجميع الطلاب، إذا كان التوزيع طبيعياً بتباين يساوى مائة؟ ما طول الفترة؟ .

الحل :

$$ن = ٢٥$$

$$\bar{س} = ٧٢$$

$$\frac{أ}{٢} = \frac{١}{٢}(١ - ٠,٩٥)$$

$$= ٠,٠٥$$

$$ى_{٠,٩٥} = ١,٦٤$$

$$م = \sqrt{١٠٠}$$

$$= ١٠$$

$$\bar{س} + ى_{٠,٩٥} \frac{م}{\sqrt{ن}} > و > \bar{س} - ى_{٠,٩٥} \frac{م}{\sqrt{ن}} \qquad (٦)$$

∴ الحد الأدنى لفترة الثقة $= \bar{س} - ى_{٠,٩٥} \frac{م}{\sqrt{ن}}$

الحد الأعلى لفترة الثقة $= \bar{س} + ى_{٠,٩٥} \frac{م}{\sqrt{ن}}$

∴ الحد الأدنى $= ٧٢ - ١,٦٤ \times \frac{١٠}{٥}$

$= ٧٢ - ٣,٢٨$

$= ٦٨,٧$

الحد الأعلى $= ٧٢ + ٣,٢٨$

$= ٧٥,٢٨$

طول الفترة = الحد الأعلى ـ الحد الأدنى

$= ٧٥,٢٨ - ٦٨,٧٢$

$= ٦,٥٦$

إذاً طول الفترة $= ٢ ى_{(١ - \frac{أ}{٢})} \frac{م}{\sqrt{ن}}$

بما أن

$$\text{ى} = \frac{\bar{\text{س}} - \text{و}}{\text{م} / \sqrt{\text{ن}}} \qquad (٣)$$

فإن :

$$\text{م ى} = (\bar{\text{س}} - \text{و}) \sqrt{\text{ن}} \qquad (٧)$$

وعليه تكون :

$$\sqrt{\text{ن}} = \frac{\text{م ى}}{(\bar{\text{س}} - \text{و})} \qquad (٨)$$

والمعادلة السابقة توضح الحجم المناسب للعينة، والذى بموجبه يمكن أن يكون الفرق بين وسط العينة ووسط المجتمع بقدر معلوم وباحتمال محدد. هذا، ويلاحظ أن حجم العينة يزداد كلما قل ذلك الفرق، أى أن زيادة حجم العينة يزيد من دقة التقدير.

البرنامج التالى يقوم بحساب وتحديد الحد الأدنى والحد الأعلى للثقة ـ بمستوى ٩٥٪ ـ وطول الفترة باستخدام المعادلة :

$$X \pm C \sqrt{\frac{V}{N}}$$

حيث :

X = الوسط الحسابى

C = القيمة الطبيعية المجدولة = ١,٦٤

V = التباين

N = حجم العينة

```
10 REM  برنامج لحساب الحد الادنى والحد الاعلى للثقد وطول الفترد
20 READ N,X,V,C
30 DATA 25,72,100,1.64
40 L=X-C*(SQR(V)/SQR(N))    REM  الحد الادنى
50 H=X+C*(SQR(V)/SQR(N))    REM  الحد الاعلى
60 I=H-L        REM  طول الفترد
70 PRINT TAB(20);L;'= الحد الادنى'
80 PRINT TAB(20);H;'= الحد الاعلى'
90 PRINT TAB(20);I;'= طول الفترد'
100 PRINT
110 END

                    المخرجات
                    ---------

          68.71999   = الحد الادنى
          75.28     = الحد الأعلى
          6.560013   = طول الفترد
```

مثال (٧،٢) :

ما هو حجم العينة الذى بموجبه يمكن التأكد ، بمستوى ٩٥٪ ، من أن تقدير وسط العينة لن يكون مخطئاً بأكثر من ٣ وحدات ، عن وسط المجتمع فى المثال السابق؟

الحل :

$$\sqrt{\text{ن}} = \frac{\text{م ى}}{(\bar{\text{س}} - \text{و})} \qquad (٨)$$

$$= \frac{١٫٦٤ \times ١٠}{٣}$$

$$\therefore \quad \text{ن} = \frac{١٠٠ \times (١٫٦٤)^{٢}}{٩}$$

$$= ٣٠$$

تكون قيمة تباين المجتمع مجهولة فى أكثر الحالات ، خاصة فى المجالات الاجتماعية ، لذلك يمكن استبدالها بتباين العينـة إذا كان حجم العينة كبيراً. أما إذا كان حجم العينة صغيراً جداً ـ أقل من ٢٠ أو ٣٠ ـ فلا مناص من استبدال التوزيع الطبيعى بتوزيع ت على عـدد (ن ـ ١) درجة حرية ، وذلك لأن :

$$\text{ت}_{(\text{ن} - ١)} = \frac{\bar{\text{س}} - \text{و}}{\text{ع} / \sqrt{\text{ن}}} \qquad (٩)$$

أما بقية المعادلة الخاصة بحدود الثقة فتظل على ما كانت عليه ، بعد استبدال القيم المعيارية بقيم ت .

مثال (٧،٣) :

أخذت عينة من درجات ٢٥ طالباً فى مادة الرياضيات ، وكان الوسط الحسابى للعينة ٧٢ درجة ، والانحراف المعيارى للعينة أيضاً ٨ درجات. فما فترة ٩٠٪ ثقة للوسط الخاص بالمجتمع؟

الحل :

$$\bar{س} + ن \frac{ع}{\sqrt{ن}} > و > \bar{س} - ت \frac{ع}{\sqrt{ن}} \quad (١٠)$$

$\sqrt{ن}$ = ٥

$\bar{س}$ = ٧٢

ع = ٨

$\frac{أ}{٢}$ = ٠,٠٥ ⟸ ١ - $\frac{أ}{٢}$ = ٠,٩٥

$(ت_{٢٤ ، ٠,٩٥})$ = ١,٧١١

$$٧٢ + ١,٧١١ \times \frac{٨}{٥} > و > ٧٢ - ١,٧١١ \times \frac{٨}{٥}$$

$$٧٤,٧ > و > ٦٩,٣$$

البرنامج التالى يقوم بتحديد الحد الأدنى والحد الأعلى لفترة الثقة بمستوى ٩٠٪ للوسط الخاص بالمجتمع باستخدام المعادلة :

$$x \pm C \sqrt{\frac{V}{N}}$$

حيث :

X = الوسط الحسابى للعينة

C = ١,٧١١ = قيمة ت المجدولة على (ن - ١)

N = حجم العينة = ن

```
10 REM     برنامج لحساب فتره الثقه
20 READ N,X,V,C
30 DATA 25,72,8,1.711
40 L=X-C*(V/SQR(N))    REM  الحد الادنى
50 H=X+C*(V/SQR(N))    REM  الحد الاعلى
60 PRINT TAB(20);L;'= الحد الادنى '
70 PRINT TAB(20);H;'= الحد الاعلى '
80 PRINT
90 END

                         المخرجات
                         ---------

                 69.26239   = الحد الادنى
                 74.73759   = الحد الاعلى
```

٣ ـ فترة الثقة للفرق بين وسطين :

اتضح من الفصل الماضى أن توزيع الفرق بين وسطى العينتين التابعتين لتوزيعين طبيعيين يكون على النحو التالى :

$$\bar{س}_1 - \bar{س}_2 \sim ط\left(و_1 - و_2 ،\ \frac{م_1^2}{ن_1} + \frac{م_2^2}{ن_2}\right) \qquad (11)$$

وعليه تكون فترة الثقة هى :

$$(\bar{س}_1 - \bar{س}_2) + ى_{\frac{أ}{2}}\, م_{21} \leqslant و_1 - و_2 \leqslant (\bar{س}_1 - \bar{س}_2) - ى_{\frac{أ}{2}}\, م_{21} \qquad (12)$$

حيث

$$م_{21} = \sqrt{\frac{م_1^2}{ن_1} + \frac{م_2^2}{ن_2}} \qquad (13)$$

مثال (٤،٧) :

كان الوسط الحسابى لرواتب عينة من العاملين بإحدى المؤسسات ٦٠٠٠ ريال، بينما كان الوسط الحسابى لعينة أخرى من العاملين فى مؤسسة ثانية يساوى ٥٣٠٠ ريال. فإذا كان التباين فى المؤسسة الأولى يساوى ٢٠٠٠ وفى الثانية ١٦٠٠ فأوجد فترة ٩٥٪ ثقة للفرق بين الوسطين، إذا كان حجم العينة الأولى ٢٥ والثانية ٤٠ شخصاً.

الحل :

$$(1 - أ) = 0.95$$

$$\therefore \frac{أ}{2} = 0.025$$

$$ى_{\frac{أ}{2}} = 1.96$$

$$(\bar{س}_1 - \bar{س}_2) + ى_{\frac{أ}{2}}\, م_{21} \leqslant و_1 - و_2 \leqslant (\bar{س}_1 - \bar{س}_2) - ى_{\frac{أ}{2}}\, م_{21} \qquad (12)$$

$$م_{21} = \sqrt{\frac{م_1^2}{ن_1} + \frac{م_2^2}{ن_2}} \qquad (13)$$

$$= \sqrt{\frac{٢٠٠٠}{٢٥} + \frac{١٦٠٠}{٤٠}}$$

$= ١٠,٩٥٤$

$\bar{س}_١ - \bar{س}_٢ = ٦٠٠٠ - ٥٣٠٠$

$= ٧٠٠$

$ى_{\frac{١}{٢}} م_{٢١} = ١,٩٦ \times ١٠,٩٥٤$

$= ٢١,٤٥$

الحد الأدنى $= ٧٠٠ - ٢١,٤٥$

$= ٦٧٨,٥٥$

الحد الأعلى $= ٧٠٠ + ٢١,٤٥$

$= ٧٢١,٤٥$

أما إذا كان التباينان مجهولين، ولكنهما متساويان، ففترة الثقة للفرق بين الوسطين هى :

$$(\bar{س}_١ - \bar{س}_٢) + ت\, ع_{٢١} \geqslant و_١ - و_٢ \geqslant (\bar{س} - س_٢) - ت\, ع_{٢١} \qquad (١٤)$$

حيث ت تعنى توزيع ت على $(ن_١ + ن_٢ - ٢)$ درجات حرية

كما أن :

$$ع^٢_{٢١} = \left(\frac{(ن_١ - ١) ع^٢_١ + (ن_٢ - ١) ع^٢_٢}{(ن_١ + ن_٢ - ٢)^٢} \right) \left(\frac{١}{ن_١} + \frac{١}{ن_٢} \right) \qquad (١٥)$$

البرنامج التالى لحساب حدود الثقة للفرق بين وسطين للبيانات بالمثال (٤ , ٧) السابق ، علماً بأن مستوى الثقة ٩٥٪ والمعادلة المستخدمة هى :

$$(\bar{X}_1 - \bar{X}_2) \pm B$$

حيث :

الفرق بين الوسطين الحسابيين للعينتين = $X_1 - X_2$

$$B = C\sqrt{\frac{V_1}{N_1} + \frac{V_2}{N_2}}$$

القيمة الطبيعية المجدولة = ١,٩٦ = C

تباين المجتمع الأول = V_1

تباين المجتمع الثانى = V_2

حجم العينة الأولى = N_1

حجم العينة الثانية = N_2

```
10 REM      برنامج لحساب فتره الثقه للفرق بين وسطين
20 READ N1,N2,X1,X2,V1,V2,C
30 DATA 25,40,6000,5300,2000,1600,1.96
40 A=SQR(V1/N1+V2/N2)
50 B=C*A
60 L=X1-X2-B    REM  الحد الادنى
70 H=X1-X2+B    REM  الحد الاعلى
80 PRINT TAB(20);L;'= الحد الادنى'
90 PRINT TAB(20);H;'= الحد الاعلى'
100 PRINT
110 END

                    المخرجات
                    --------

           678.5291   = الحد الادنى
           721.4707   = الحد الاعلى
```

مثال (٧,٥) :

افرض أن التباينين غير معلومين فى المثال السابق، وافرض أن الانحراف المعيارى للعينة الأولى يساوى ١٠، بينما كان الانحراف المعيارى للعينة الثانية ٨. أوجد فترة ٩٥٪ ثقة للفرق بين الوسطين، بافتراض أن التباينين متساويان فى المجتمعين.

الحل :

$$ن_١ = ٢٥$$

$$ن_٢ = ٤٠$$

$$ن_١ + ن_٢ - ٢ = ٦٣$$

$$ت = ٢,٠$$

$$ع^٢_{٢١} = \left(\frac{٢٤ \times ١٠٠ + ٣٩ \times ٦٤}{(٢٥ + ٤٠ - ٢)}\right)\left(\frac{١}{٢٥} + \frac{١}{٤٠}\right) \qquad (١٥)$$

$$= ٥,٠٥$$

$$ع_{٢١} = ٢,٢٥$$

$$\bar{س}_١ - \bar{س}_٢ = ٧٠٠$$

$$٧٠٠ + ٢ \times ٢,٢٥ \leqslant و_١ - و_٢ \leqslant ٧٠٠ - ٢ \times ٢,٢٥$$

$$٧٠٤,٥ \leqslant و_١ - و_٢ \leqslant ٦٩٥,٥$$

البرنامج التالى لحساب حدود الثقة بمستوى معنوية ٩٥٪ للبيانات الواردة للفرق بين وسطين فى حالة تساوى التباينين للمجتمعين مع عدم معرفتهما، وذلك باستخدام المعادلة :

$$(X_1 - X_2) \pm B$$

حيث :

$\underline{X_1}$ = وسط العينة الأولى

$\underline{X_2}$ = وسط العينة الثانية

$B = CA$

C = ٢,٠ = قيمة ت المجدولة على (ن ١ + ن ٢ - ٢)

$$A = \left[\frac{(N_1 - 1)V_1 + (N_2 - 1)V_2}{N_1 + N_2 - 2}\right]\left[\frac{1}{N_1} + \frac{1}{N_2}\right]$$

N_1 = حجم العينة الأولى

N_2 = حجم العينة الثانية

V_1 = تباين العينة الأولى

V_2 = تباين العينة الثانية

```
10  REM  برنامج لحساب حدود الثقه للفرق بين وسطين فى حاله
20  REM  تساوى التباينين للمجتمعين مع عدم معرفتهما
30  READ N1,N2,X1,X2,V1,V2,C
40  DATA 25,40,6000,5300,10,8,2
50  V1=V1*V1
60  V2=V2*V2
70  A=(((N1-1)*V1+(N2-1)*V2)/(N1+N2-2))*(1/N1+1/N2)
80  X=ABS(X1-X2)
90  B=C*SQR(A)
100 L=X-B              REM  الحد الادنى
110 H=X+B              REM  الحد الاعلى
120 PRINT TAB(20);L;'= الحد الادنى'
130 PRINT
140 PRINT TAB(20);H;'= الحد الاعلى'
150 PRINT
999 END
```

المخرجات

```
695.5049   = الحد الادنى
704.4949   = الحد الاعلى
```

٤ ـ حدود الثقة للنسب :

جاء فى الفصـل السابق أنه إذا كانت $\hat{ح}$ تعنى نسبة أفراد العينة الذين يتميزون بصفة معينة، فإن :

$$\frac{\hat{ح} - ح}{\sqrt{\frac{ح(١-ح)}{ن}}} \sim ط(٠، ١) \qquad (١٦)$$

حيث ح هى نسبة أفراد المجتمع الذى سحبت منه تلك العينة. وبالمقارنة بحدود الثقة لوسط المجتمع تكون :

$$\hat{ح} - ى_{١-\frac{أ}{٢}} \sqrt{\frac{ح(١-ح)}{ن}} \leq ح \leq \hat{ح} + ى_{١-\frac{أ}{٢}} \sqrt{\frac{ح(١-ح)}{ن}} \qquad (١٧)$$

مثال (٦,٧) :

أجـريت دراسـة لتقدير عدد الموظفين الذين يوافقون على نظام جديد للدوام الرسمى. سحبت عينة عشوائية حجمها مائة شخص من بين العدد الكلى للموظفين والبالغ ٢٦٠٠٠ موظف، فأجاب ٣٥ موظفاً بالموافقة على النظام الجديد. فما هى فترة الثقة بمستوى ٩٥٪ لنسبة الموافقين، وبكم تقدر عدد الموافقين، وما هو الحجم المناسب للعينة، حتى لا يختلف تقدير نسبة العينة عن المجتمع بأكثر من ٨٪؟

الحل :

$$ن = ١٠٠$$

$$\hat{ح} = \frac{\text{عدد الموافقين}}{ن}$$

$$\therefore \hat{ح} = ٠,٣٥$$

$$١ - \hat{ح} = ٠,٦٥$$

$$ى_{١-\frac{أ}{٢}} = ى_{٠,٩٧٥}$$

$$= ١,٩٦$$

إذاً فترة الثقة هى :

$$\text{٠,٣٥} - \text{١,٩٦} \times \sqrt{\frac{\text{٠,٣٥} \times \text{٠,٦٥}}{\text{١٠٠}}} < \text{ح} < \text{٠,٣٥} + \text{١,٩٦} \times \sqrt{\frac{\text{٠,٣٥} \times \text{٠,٦٥}}{\text{١٠٠}}}$$

$$\text{٠,٢٥} \leqslant \text{ح} \leqslant \text{٠,٤٤}$$

والعدد الكلى يتراوح بين :

$$\text{٢٦٠٠٠٠} \times \text{٠,٢٥} \leqslant \text{العدد الكلى} \leqslant \text{٢٦٠٠٠٠} \times \text{٠,٤٤}$$

$$\text{٦٥٠٠٠} \leqslant \text{العدد الكلى} \leqslant \text{١١٤٤٠٠}$$

أما حجم العينة المناسب فيمكن استخراجه من المعادلة

$$\text{ح} - \acute{\text{ح}} = \text{ى}_{\text{١}-\frac{\text{ا}}{\text{٢}}} \sqrt{\frac{\acute{\text{ح}}(\text{١}-\acute{\text{ح}})}{\text{ن}}} \qquad (\text{١٨})$$

$$(\text{ح} - \acute{\text{ح}})^{\text{٢}} = \text{ى}^{\text{٢}}_{\text{١}-\frac{\text{ا}}{\text{٢}}} \left(\frac{\acute{\text{ح}}(\text{١}-\acute{\text{ح}})}{\text{ن}}\right) \qquad (\text{١٩})$$

ومن ثم :

$$\text{ن} = \text{ى}^{\text{٢}}_{\text{١}-\frac{\text{ا}}{\text{٢}}} \frac{\acute{\text{ح}}(\text{١}-\acute{\text{ح}})}{(\text{ح}-\acute{\text{ح}})^{\text{٢}}} \qquad (\text{٢٠})$$

$$\text{ن} = (\text{١,٩٦})^{\text{٢}} \times \frac{\text{٠,٣٥} \times \text{٠,٦٥}}{(\text{٠,٠٨})^{\text{٢}}}$$

= ١٣٧ شخصاً.

إذاً يجب زيادة حجم العينة السابقة بسبعة وثلاثين شخصاً؛ للحصول على الدقة المطلوبة . يقترب التوزيع ذو الحدين من التوزيع الطبيعى كلما ازداد حجم العينة ، إلا أنه لا يصبح التوزيع الطبيعى الأمثل إلا إذا كان حجم العينة كبيراً جداً. لذلك تختلف الحدود الدقيقة بعض الشيء عن المقدرة بالمعادلة السابقة ، ولكن عرض الفترة يظل سليماً. ولكثرة استخدامات فترات الثقة فى المجالات الاجتماعية ، فقد دونت الحدود الدقيقة لبعض العينات . والجدول التالى يوضح تلك الحدود .

جدول (١)

فترات الثقة للنسب باستخدام ذى الحدين[١].

95% CONFIDENCE INTERVAL (PER CENT) FOR BINOMIAL DISTRIBUTION (1)*

Number Observed	Size of Sample, n												Fraction Observed	Size of Sample			
f	10		15		20		30		50		100		f/n	250		1000	
0	0	27	0	20	0	15	0	10	0	07	0	4	0.00	0	1	0	0
1	0	40	0	31	0	23	0	17	0	11	0	5	.01	0	4	0	2
2	3	61	2	37	1	30	1	21	0	14	0	7	02	1	5	1	3
3	8	62	5	45	4	36	2	25	1	17	1	8	03	1	6	2	4
4	15	74	9	56	7	42	4	30	2	19	1	10	.04	2	7	3	5
5	22	78	14	64	10	47	6	33	3	22	2	11	.05	3	9	4	7
6	26	85	19	67	14	54	9	37	5	24	2	12	.06	3	10	5	8
7	38	92	19	71	14	59	10	41	6	27	3	14	.07	4	11	6	9
8	39	97	29	81	20	65	13	44	7	29	4	15	.08	5	12	6	10
9	60	100	33	81	22	71	16	48	9	31	4	16	.09	6	13	7	11
10	73	100	36	86	29	71	17	53	10	34	5	18	.10	7	14	8	12
11			44	91	29	78	20	56	12	36	5	19	11	7	16	9	13
12			55	95	35	80	23	60	13	38	6	20	12	8	17	10	14
13			63	98	41	86	24	64	15	41	7	21	.13	9	18	11	15
14			69	100	47	86	29	68	16	43	8	22	.14	10	19	12	16
15			80	100	53	90	32	68	18	44	9	24	.15	10	20	13	17
16					58	93	32	71	20	46	9	25	.16	11	21	14	18
17					64	96	36	76	21	48	10	26	17	12	22	15	19
18					70	99	40	77	23	50	11	27	.18	13	23	16	21
19					77	100	44	80	25	53	12	28	.19	14	24	17	22
20					85	100	47	83	27	55	13	29	.20	15	26	18	23
21							52	84	28	57	14	30	.21	16	27	19	24
22							56	87	30	59	14	31	.22	17	28	19	25
23							59	90	32	61	15	32	.23	18	29	20	26
24							63	91	34	63	16	33	.24	19	30	21	27
25							67	94	36	64	17	35	25	20	31	22	28
26							70	96	37	66	18	36	26	20	32	23	29
27							75	98	39	68	19	37	.27	21	33	24	30
28							79	99	41	70	19	38	.28	22	34	25	31
29							83	100	43	72	20	39	.29	23	35	26	32
30							90	100	45	73	21	40	.30	24	36	27	33
31									47	75	22	41	.31	25	37	28	34
32									50	77	23	42	.32	26	38	29	35
33									52	79	24	43	.33	27	39	30	36
34									54	80	25	44	34	28	40	31	37
35									56	82	26	45	.35	29	41	32	38
36									57	84	27	46	.36	30	42	33	39
37									59	85	28	47	.37	31	43	34	40
38									62	87	28	48	.38	32	44	35	41
39									64	88	29	49	39	33	45	36	42
40									66	90	30	50	.40	34	46	37	43
41									69	91	31	51	.41	35	47	38	44
42									71	93	32	52	.42	36	48	39	45
43									73	94	33	53	43	37	49	40	46
44									76	95	34	54	.44	38	50	41	47
45									78	97	35	55	.45	39	51	42	48
46									81	98	36	56	.46	40	52	43	49
47									83	99	37	57	47	41	53	44	50
48									86	100	38	58	.48	42	54	45	51
49									89	100	39	59	.49	43	55	46	52
50									93	100	40	60	.50	44	56	47	53
										†		†			††		††

† If f exceeds 50, read $100 - f$ = number observed and subtract each confidence limit from 100.

†† If f/n exceeds 0.50, read $1.00 - f/n$ = fraction observed and subtract each confidence limit from 100.

(١) المصدر :

Snedecor (G. W.) and Cochran (W.G.); Statistical Methods, Iwoa University Press, Iwoa, U.S.A.; Seventh Printing, Sixth edition; Page (6).

تابع جدول (١)

99% Confidence Interval (Per Cent) for Binomial Distribution (1)*

Number Observed f	Size of Sample, n: 10		15		20		30		50		100		Fraction Observed f/n	Size of Sample: 250		1000	
0	0	38	0	28	0	21	0	16	0	10	0	5	0.00	0	2	0	1
1	0	52	0	38	0	30	0	21	0	14	0	7	01	0	5	0	2
2	1	63	1	47	0	38	0	26	0	17	0	9	02	1	6	1	3
3	4	71	3	54	2	43	1	31	1	20	0	10	03	1	7	2	4
4	9	79	5	63	4	50	2	35	1	23	1	12	04	2	9	3	6
5	15	85	9	68	6	58	4	39	2	26	1	13	05	2	10	3	7
6	21	91	13	73	9	61	6	43	3	29	2	14	06	3	11	4	8
7	29	96	17	78	12	64	8	47	4	31	2	16	07	3	13	5	9
8	37	99	22	83	16	71	10	51	6	33	3	17	08	4	14	6	10
9	48	100	27	87	20	73	12	54	7	36	3	18	09	5	15	7	12
10	62	100	32	91	20	80	15	57	8	38	4	19	10	6	16	8	13
11			37	95	27	80	15	62	10	40	4	20	11	6	17	9	14
12			46	97	29	84	19	66	11	43	5	21	12	7	18	9	15
13			53	99	36	88	20	68	12	45	6	23	13	8	19	10	16
14			62	100	39	91	24	70	14	47	6	24	14	9	20	11	17
15			72	100	42	94	25	75	15	49	7	26	15	9	22	12	18
16					50	96	30	76	17	51	8	27	16	10	23	13	19
17					57	98	32	80	18	53	9	29	.17	11	24	14	20
18					62	100	34	81	20	55	9	30	.18	12	25	15	21
19					70	100	38	85	21	57	10	31	19	13	26	16	22
20					79	100	43	85	23	59	11	32	20	14	27	17	23
21							46	88	24	61	12	33	21	15	28	18	24
22							49	90	26	63	12	34	22	16	30	19	26
23							53	92	28	65	13	35	.23	17	31	20	27
24							57	94	29	67	14	36	24	18	32	21	28
25							61	96	31	69	15	38	.25	18	33	22	29
26							65	98	33	71	16	39	.26	19	34	22	30
27							69	99	35	72	16	40	.27	20	35	23	31
28							74	100	37	74	17	41	28	21	36	24	32
29							79	100	39	76	18	42	29	22	37	25	33
30							84	100	41	77	19	43	30	23	38	26	34
31									43	79	20	44	31	24	39	27	35
32									45	80	21	45	32	25	40	28	36
33									47	82	21	46	33	26	41	29	37
34									49	83	22	47	34	26	42	30	38
35									51	85	23	48	.35	27	43	31	39
36									53	85	24	49	36	28	44	32	40
37									55	88	25	50	37	29	45	33	41
38									57	89	26	51	38	30	46	34	42
39									60	90	27	52	.39	31	47	35	43
40									62	92	28	53	40	32	48	36	44
41									64	93	29	54	41	33	50	37	45
42									67	94	29	55	42	34	51	38	46
43									69	96	30	56	43	35	52	39	47
44									71	97	31	57	44	36	53	40	48
45									74	98	32	58	45	37	54	41	49
46									77	99	33	59	46	38	55	42	50
47									80	99	34	60	.47	39	55	43	51
48									83	100	35	61	48	40	56	44	52
49									86	100	36	62	49	41	57	45	53
50									90	100	37	63	50	42	58	46	54
												†			††		††

† If f exceeds 50, read $100 - f$ = number observed and subtract each confidence limit from 100

†† If f/n exceeds 0.50, read $1.00 - f/n$ = fraction observed and subtract each confidence limit from 100

٥ ـ فترات الثقة للتباينات :

بما أن :

$$\frac{(ن - ١)ع^{٢}}{م^{٢}} \sim ك^{٢}_{(ن - ١)} \qquad (٢١)$$

وبما أن توزيع مربع كاى ($ك^{٢}$) غير متشابه، ففترة الثقة بمستوى (١ ـ أ)٪ هى :

$$ك^{٢}_{\frac{أ}{٢}} < \frac{(ن - ١)ع^{٢}}{م^{٢}} \leqslant ك^{٢}_{(١ - \frac{أ}{٢})} \qquad (٢٢)$$

إذاً :

$$\frac{(ن - ١)ع^{٢}}{ك^{٢}_{١ - \frac{أ}{٢}}} < م^{٢} < \frac{(ن - ١)ع^{٢}}{ك^{٢}_{\frac{أ}{٢}}} \qquad (٢٣)$$

مثال (٧,٧) :

أخذت عينة عشوائية حجمها ١٢ من مجتمع طبيعى ، فوجد أن تباينها يساوى ٥٣ . أوجد فترة ٩٥٪ ثقة للتباين.

الحل :

$١ - أ = ٠,٩٥$

$\therefore أ = ٠,٠٥$

$\frac{أ}{٢} = ٠,٠٢٥$

$ن = ١٢$

$ن - ١ = ١١$

$ك^{٢}_{١ - \frac{أ}{٢}} = ك^{٢}_{٠,٩٧٥}$ على ١١ درجة حرية.

$= ٢١,٩$

$ك^{٢}_{٠,٠٢٥}$ على ١١ درجة حرية $= ٣,٨٢$

$\therefore$ فترة الثقة هى :

$$\frac{١١ \times ٥٣}{٢١,٩} < م^{٢} < \frac{١١ \times ٥٣}{٣,٨٢} \qquad (٢٣)$$

$$٢٦,٦٢ < م^{٢} < ١٥٢,٦٢$$

تمارين

١ ـ عرف فترة الثقة، واستخداماتها، والفرق بينها، وبين حدود الثقة.

٢ ـ ما هو الفرق بين فترة الثقة واختبار الفرضية؟ وهل يجوز أن تكون فترة الثقة بديلاً لاختبار الفرضية فى حالة خاصة؟

٣ ـ ترغب إحدى المؤسسات فى شراء مصابيح كهربائية من نوع خاص، فعرضت عليها ثلاث شركات أنواعاً مختلفة من تلك المصابيح. وباختيار عينة عشوائية حجمها مائة مصباح من كل نوع اتضح أن الوسط الحسابى لعدد أيام الإنارة المستمرة لعينة كل نوع، وتباين المجتمع على النحو الآتى :

الوسط للنوع الأول ٤٣,٤ يوماً، والتباين ٣٤,٨١ للمجتمع الأول.
الوسط للنوع الثانى ٥١,٣ يوماً، والتباين ٤٦,٢٤ للمجتمع الثانى.
الوسط للنوع الثالث ٤٤,٧ يوماً، والتباين ٥٣,٢٩ للمجتمع الثالث.
استخدم ٩٥٪ فترة ثقة لوسط كل نوع، وحدد أى الأنواع أفضل.

٤ ـ استخدم البيانات الخاصة بالسؤال الثالث لإيجاد فترة ٩٥٪ ثقة للفرق بين الوسطين لكل عينتين، ومن ثم قرر ما إذا كان هناك نوعان متساويان أم لا.

٥ ـ افرض أن حجم العينة للنوع الأول فى السؤال الثالث ١٥، وللنوع الثانى ٢٠، وللنوع الثالث ١٢، فأوجد فترة ٩٥٪ ثقة للفرق بين كل وسطين، إذا علمت أن التجارب السابقة قد دلت على أن التباين لا يختلف بين المجتمعات الثلاثة، بينما كانت الانحرافات المعيارية للعينات ٦ و ٨ و ٧ على التوالى.

٦ ـ أوجد فترة ٩٥٪ ثقة لأوساط المصابيح الثلاثة الواردة فى السؤال الخامس.

٧ ـ أوجد فترة ٩٥٪ ثقة للفرق بين كل وسطين فى السؤال الخامس، إذا لم تكن هناك أى معلومات متوفرة حول تباين المجتمع لكل نوع.

٨ ـ تعاطى ١٠ مرضى بالسرطان، اختيروا كعينة عشوائية، دواءً جديداً يساعد على زيادة عدد ساعات النوم، فكانت الزيادات فى أحد الأيام على النحو الآتى :

٣,٥ ، ٦,٣ ، ٢,١ ، ١,٦ ، ٠,٨ ، ٤,٧ ، ٤,١ ، ٠,٠ ، ٥,٤ ، ٢,٥

فأوجد فترة ٩٠٪ ثقة لوسط المجتمع.

٩ ـ أوجد فترة ٩٩٪ ثقة للمجتمع إذا دلت التجارب السابقة على أن تباين المجتمع ٦,٢٥ .

١٠ ـ أوجد حجم العينة المناسب الذى يمكن أن يستخدم بمستوى ٩٥٪ ثقة، لتقدير الوسط للعدد اليومى للمعاملات فى حدود ± ٤ معاملات، إذا علمت أن الانحراف المعيارى للعدد اليومى للمعاملات يساوى ١٥ معاملة.

١١ ـ افرض أن الانحراف المعيارى للمجتمع فى المثال السابق لم يكن معلوماً، ولكن باختيار عينة من معاملات ١٠ أيام اتضح أن الانحراف المعيارى لتلك العينة يساوى ١٢، فما الحجم المناسب للعينة؟

١٢ ـ اختيرت عينة عشوائية قوامها ٧٥ وحدة من إنتاج أحد المصانع للفحص، فاتضح أن ١٤ وحدة كانت تالفة. أوجد فترة ٩٥٪ فترة ثقة لنسبة الوحدات التالفة، وبكم تقدر عدد الوحدات التالفة يومياً، إذا كان إنتاج المصنع ١٠٠٠٠ وحدة فى اليوم.

١٣ ـ ما هو الحجم المناسب للعينة فى السؤال السابق، إذا كان الهدف هو تقدير نسبة التالف بنسبة تختلف ٥٪ على الأكثر من نسبة المجتمع.

١٤ ـ أوجد فترة ٩٠٪ ثقة لتباين مجتمع طبيعى، إذا علمت أن تباين عينة عشوائية قوامها ٢٤ يساوى ٦٠.

١٥ ـ استخدم بيانات السؤال الثالث واكتب برنامجاً بلغة بيسك لتحديد الحد الأدنى والحد الأعلى للثقة.



تطبيقات اختبارات الفرضيات
(Hypothesis Testing)



تطبيقات اختبارات الفرضيات
(Hypothesis Testing)

١ - تعريف الفرضية والاختبار :

الفرضيات كلمة جمع، مفردها فرضية، والفرضية هى بيان يتعلق بالتوزيع الإحصائى للمتغير العشوائى ؛ لذلك تنقسم الفرضيات إلى نوعين هما :

١ - فرضية تتعلق بمعلم واحد، أو أكثر من معالم (Parameters) التوزيع، الذى يفترض أنه معلوم . فقد تكون الفرضية مثلا خاصة بنسبة معينة لقيم تتبع التوزيع ذا الحدين، أو قد تكون خاصة بالوسط أو الفرق بين وسطين لقيم تتبع التوزيع الطبيعى على سبيل المثال أيضا . . . وهكذا.

٢ - فرضية تتعلق بالتوزيع الإحصائى نفسه مثل تبعية متغير عشوائى لتوزيع معين.

هذا وسيعرض هنا النوع الأول فقط لكثرة استخداماته فى المجالات التطبيقية. لذلك فسوف يفترض أن التوزيع الإحصائى للقيم معلوم، لتصبح الفرضية معتمدة على القيم العينية فى اتخاذ القرار الخاص بقبول أو رفض بيان أو عدة بيانات تتعلق ببعض معالم ذلك التوزيع.

تسمى الفرضية بالفرضية البسيطة (Simple Hypothesis) إذا حددت قيمة معينة للمعلم، وتسمى مزدوجة (Composite Hypothesis) إذا كانت بخلاف ذلك.

مثال (٨،١) :

س متغير عشوائى يتبع توزيعاً طبيعياً بتباين = ١٠ .

(أ) الفرضية القائلة بأن الوسط = ٢٠ هى فرضية بسيطة.

(ب) الفرضية القائلة بأن الوسط < ٢٠ هى فرضية مزدوجة.

(جـ) الفرضية القائلة بأن الوسط ينحصر فى الفترة من ١٥ إلى ٢٠ هى فرضية مزدوجة أيضاً.

أما اختبار الفرضية فهو عبارة عن تقسيم فضاء العينة الذى يحوى كل النتائج المتوقعة إلى قسمين منفصلين، الأول : يتكون من جميع النتائج التى تدعو لقبول الفرضية، والثانى : يتكون من جميع النتائج الداعية لرفض تلك الفرضية.

لذا فاختبار الفرضية هو أسلوب لاتخاذ أحد قرارين : إما القبول أو الرفض، بناء على تقسيم فضاء العينة إلى منطقتين غير متداخلتين، أولاهما تسمى منطقة القبول (Acceptance Region) ، والثانية تسمى منطقة الرفض (Rejection Region) أو المنطقة الحرجة (Critical Region) .

بيد أن اتخاذ القرار يتم تحت ظروف عدم التأكد، وهذا يعنى أن هناك احتمالاً برفض الفرضية الصحيحة، أو قبول الفرضية الخاطئة، بمعنى أن هناك أربعة قرارات لا بد من أن يتخذ واحد منها، وهى :

١ ـ قبول الفرضية عندما كان يجب أن تقبل، وهو قرار صحيح .
٢ ـ رفض الفرضية عندما كان يجب أن تقبل، وهو قرار خاطىء.
٣ ـ قبول الفرضية عندما كان يجب أن ترفض، وهو قرار خاطىء.
٤ ـ رفض الفرضية عندما كان يجب أن ترفض، وهو قرار صحيح .

ومن الواضح أن النوع الأول والثانى يمثلان مجموعة قرارات تنفصل تماماً عن الثالث والرابع، وهذا دليل على أن اختبار الفرضيات هو أسلوب لاتخاذ قرارين: أحدهما صحيح، مثل الأول أو الرابع، والثانى: خاطىء مثل القرار الثانى والثالث. ويقال إن هناك خطأ من النوع الأول (Type I error) ، إذا رفضت الفرضية عندما كانت صحيحة. أما إذا قبلت الفرضية الخاطئة فيسمى الخطأ من النوع الثانى (Type II error) . والجدول التالى يبين أنواع القرارات والأخطاء، علماً بأن :

أ = احتمال ارتكاب خطأ من النوع الأول.
ب = احتمال ارتكاب خطأ من النوع الثانى.

وعليه يصبح احتمال اتخاذ القرار الأول والخاص بقبول الفرضية عندما كان يجب أن تقبل يساوى (١ ـ أ). أما احتمال اتخاذ القرار الخاص برفض الفرضية عندما كان يجب أن ترفض (الرابع) فيساوى (١ ـ ب).

جدول (١)

أنواع القرارات واحتمالاتها

القرار	أنواع الفرضيات	
	الفرضية صحيحة	الفرضية خاطئة
قبول الفرضية	قرار سليم (١ - أ) مستوى الثقة	خطأ من النوع الثانى (ب)
رفض الفرضية	خطأ من النوع الأول (أ) (مستوى المعنوية)	قرار سليم (١ - ب) (قوة الاختبار)

يعتبر الخطأ من النوع الأول أكثر خطراً من النوع الثانى، لذلك تسمى (أ) بكمية المخاطرة التى يجب أن توضع فى الاعتبار عند صياغة أى فرضية.

مثال (٢, ٨) :

أنتجت شركـة للأدوية دواءً جديداً، ولا بد من التأكد من أن الدواء ليست له أعراض جانبية تؤذى الإنسان.

إذاً هناك نوعان من الفرضيات :

الفرضية الأولى : الدواء غير مؤذٍ.

الفرضية الثانية : الدواء مؤذٍ.

فأى الفرضيتين يجب أن تختار؟

الحل :

لنفرض أن الاختيار وقع على الفرضية الأولى، فالمخاطرة تأتى هنا (خطأ النوع الأول) إذا رفضت الفرضية وهى صحيحة، أى إذا اعتبر الدواء مؤذياً فى حين أنه غير مؤذٍ.

أما إذا كانت الفرضية هى الثانية فالخطأ الأول هو رفضها وهى صحيحة، أى اعتبار الدواء غير موذٍ فى حين أنه موذٍ.

وبما أن اعتبار الدواء غير مؤذٍ فى حين أنه مؤذٍ، أكثر خطراً من اعتبار الدواء مؤذياً فى حين أنه غير مؤذٍ، فالفرضية الثانية هى الصحيحة فى هذه الحالة.

وبما أن احتمال الخطأ من النوع الأول (أ) يتناقص كلما ضعفت كمية المخاطرة التى ينطوى عليها القرار الخاص بقبول الفرضية الصحيحة، فإن أ تسمى مستوى المعنوية (Level of Significance) . وأما احتمال قبول الفرضية الصحيحة (١ - أ) فيسمى مستوى الثقة (Level of Confidence) ، بينما يسمى الاحتمال الثانى (١ - ب) للقرار السليم، والخاص برفض الفرضية عندما كان يجب أن ترفض، بقوة الاختبار (Power of Test) .

يعتمد اختبار الفرضية على ثلاث قيم، هى : أ، و ب، وحجم العينة (ن). وبالرغم من أنه لا يمكن ارتكاب الخطأين فى اختبار واحد، فإن الثلاث قيم مترابطة فيما بينها، إلى حد يجعل فى الإمكان استخراج القيمة الثالثة من أى قيمتين. ولعل القرار الأمثل هو الذى تكون عنده أ = ب = صفراً.

إلا أن ذلك ليس ممكناً ما دام اتخاذ القرار يتم تحت ظروف عدم التأكد؛ لذلك يهدف اختبار الفرضية إلى اختيار الاختبار المناسب الذى يؤدى إلى إضعاف قيمة (ب)، بعد تحديد احتمال المخاطرة الذى يجعل مستوى المعنوية (أ) فى أدنى درجة ممكنة، وبمعنى آخر : اختبار الفرضيات هو أسلوب لرفع قوة الاختبار إلى أعلى درجة ممكنة مع أدنى درجة من مستوى المعنوية. هذا، وتعتبر أكثر المستويات المعنوية استخداماً هى ١٪ و ٥٪ و ١٠٪، وعليه فالفرضية دائماً صحيحة ما لم يثبت خلاف ذلك.

تسمى الفرضية فى جميع الحالات السابقة بفرضية العدم (ف.) - Null Hypothesis(Ho) - فإذا ثبت أن فرضية العدم (ف.) غير صحيحة فلا بد من فرضية بديلة ($ف_1$) - (Alternative Hypothesis(H1) - فالفرضية البديلة ($ف_1$) هى عبارة عن بيان لمنطقة الرفض، ولا يتم اختبارها فى نفس الدراسة مع أنها تستخدم لاختبار فرضية العدم.

وتنقسم الفرضيات البديلة إلى قسمين، القسم الأول : هو الفرضية البديلة بطرف واحد (One – tailed Test) وهنا تكون الفرضية البديلة إما أكبر أو أصغر من مقدار معين، فهى ذات اتجاه واحد. أما القسم الثانى : فهو الفرضية البديلة ذات الطرفين (Two – tailed Test) ، إذ تكون $ف_1$ ذات اتجاهين، الأول هو أكبر، والثانى أصغر من القيمة المحددة

مثال (٣,٨) :

إذا كانت

$$\text{ف}_{\cdot} = \frac{١}{٢}$$

فحدد أنواع الفرضيات البديلة التالية .

(أ) $\text{ف}_{١} > \frac{١}{٢}$

(ب) $\text{ف}_{١} < \frac{١}{٢}$

(جـ) $\text{ف}_{١} \neq \frac{١}{٢}$

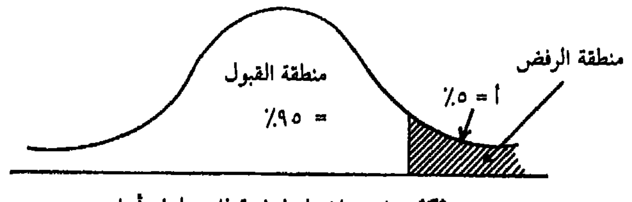
الحل :

(أ) $\text{ف}_{١} > \frac{١}{٢}$ هى فرضية ذات طرف واحد وهو الطرف الأعلى .

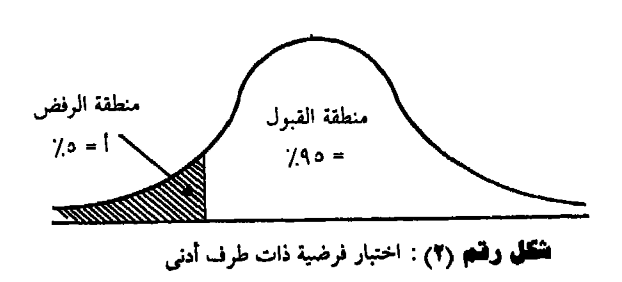
(ب) $\text{ف}_{١} < \frac{١}{٢}$ هى فرضية ذات طرف واحد وهو الطرف الأدنى .

(جـ) $\text{ف}_{١} \neq \frac{١}{٢}$ فرضية بديلة ذات طرفين .

يعتبر تحديد نوع الفرضية البديلة من الأسس التى يرتكز عليها أسلوب اختبار الفرضيات ، إذ يكون مستوى المعنوية فى اتجاه واحد ، إذا كانت الفرضية ذات اتجاه واحد ، وينقسم إلى قسمين متساويين إذا كانت الفرضية البديلة ذات اتجاهين . والأشكال التالية توضح بعض الأمثلة عنـد اختبار الفرضية لمتغير يتبع التوزيع الطبيعى ، وبمستوى معنوية يساوى ٥٪ . ويلاحظ أن منطقة القبول تساوى ٩٥٪ فى جميع الحالات .



شكل (١) : اختبار فرضية ذات طرف أعلى



شكل رقم (٢) : اختبار فرضية ذات طرف أدنى

٢ - القرار :

يعتمد القرار الخاص بقبول أو رفض فرضية العدم على ثلاث قيم، وهى :

(أ) مستوى المعنوية (أ)، أو مستوى الثقة (١ - أ) :

إذ تعزى الاختلافات الواردة بين القيم إلى الصدفة (Chance) فى حالة قبول فرضية العدم.

(ب) إحصائية الاختبار (Test Statistic) :

وهى القيمــة المحسوبة بناء على توزيع الإحصائية، فلكل إحصائية توزيع، وهى قيمة محسوبة من القيم العينية وتستخدم فى ذات الوقت القيمة المحددة بفرضية العدم. فإحصائية اختبار الوسط فى التوزيع الطبيعى هى ى. حيث :

$$\text{ى} = \frac{\bar{\text{س}} - \text{و}}{\text{م} / \sqrt{\text{ن}}} \qquad (1)$$

(جـ) القيمة الحرجة (Critical Value) :

وهى القيمة المستخرجة من جداول التوزيعات الإحصائية؛ ولذلك تسمى أيضاً بالقيمة المجدولة (Tabulated Value). وتستخرج القيمة المجدولة من جداول التوزيع الطبيعى إذا كانت إحصائية الاختبار على التوزيع الطبيعى بينما تستخرج القيمة الحرجة من توزيع ت أو ف أو مربـع كاى، وبـالعـدد المحدد لدرجات الحرية اعتماداً على توزيع إحصائية الاختبار ومستوى المعنوية.

تتساوى إحصائية الاختبار بالقيمة الحرجة عند الحد الفاصل بين منطقتى القبول والرفض؛ لذلك تقبل فرضية العدم إذا كانت القيمة المطلقة لإحصائية الاختبار أكبر من القيمة الحرجة (المجدولة). أى إذا كانت القيمة المحسوبة أكبر من المجدولة فى حالة الاختبار من الطرف الأعلى، أو أقل منها فى حالة الاختبار من الطرف الأدنى. وهذا يعنى أن فرضية العدم مقبولة

ما دامت القيمة المحسوبة واقعة ضمن فترة الثقة بمستوى معنوية محدد. هذا، وتظل فرضية العدم مقبولة إلا إذا ثبت خلاف ذلك.

مثال (٤،٨):

إذا كانت سَ هى الوسط الحسابى لعينة عشوائية حجمها (ن) من مجتمع ذى توزيع طبيعى وسطه (و) وانحرافه المعيارى (م). وإذا كانت وِ قيمة محددة ومعلومة فأوجد القيم الحرجة، وإحصائية الاختبار المرافقة لكل من الفرضيات البديلة التالية، ووضح كيفية اتخاذ القرار بمستوى معنوية يساوى (أ) إذا كانت فرضية العدم هى :

ف$_{.}$: و = وِ

والفرضيات البديلة هى :

(أ) ف$_{١}$: و ≠ وِ

(ب) ف$_{١}$: و > وِ

(جـ) ف$_{١}$: و < وِ

الحل :

ف$_{.}$: و = وِ تعنى أن وسط المجتمع الذى سحبت منه العينة يساوى مقداراً محدداً هو وِ. وقد تكتب نفس الفرضية على النحو التالى :

ف$_{.}$: و - وِ = صفراً

إحصائية الاختبار فى جميع الحالات هى :

$$ى = \frac{سَ - وِ}{م / \sqrt{ن}}$$

(أ) ترفض فرضية العدم (ف$_{.}$) إذا كانت :

$$|ى| > ى_{١ - \frac{أ}{٢}}$$

أى أن فرضية العدم مرفوضة فى حالتين هما :

$\text{ى} > \text{ى}_{1 - \frac{\text{أ}}{2}}$ إذا كانت ى > صفر

$\text{ى} < \text{ى}_{1 - \frac{\text{أ}}{2}}$ إذا كانت ى < صفر

حيث $\text{ى}_{1 - \frac{\text{أ}}{2}}$ هى القيمة الحرجة (المجدولة) بمستوى معنوية $\frac{\text{أ}}{2}$ لأن فرضية العدم ذات اتجاهين.

(ب) $\text{ف}_1 : \text{و} > \text{و}$

إذاً ترفض فرضية العدم إذا كانت

$\text{ى} > \text{ى}_{1 - \text{أ}}$

لأن الفرضية ذات طرف واحد وهو الطرف الأعلى.

(جـ) $\text{ف}_1 : \text{و} < \text{و}$

ترفض فرضية العدم إذا كانت

$\text{ى} < -\text{ى}_{1 - \text{أ}}$

لأن الفرضية ذات طرف أدنى فقط.

٢ - اختبارات الوسط الحسابى لعينة واحدة :

(أ) عندما يكون تباين المجتمع الذى سحبت منه العينة معلوماً :

إذا كانت و قيمة محددة لوسط المجتمع، بينما كانت سَ هى الوسط الحسابى لعينة حجمها (ن) اختيرت عشوائياً من مجتمع طبيعى، أثبتت التجارب السابقة أن وسطه يساوى (و) وانحرافه المعيارى م. فإحصائية الاختبار هى :

$$\text{ى} = \frac{\bar{\text{س}} - \text{و}}{\text{م} / \sqrt{\text{ن}}}$$

مثال (٥,٨):

اختيرت عينة عشوائية قوامها ٢٥ خريجاً من أحد برامج النسخ، اتضح بعد اختبار أفراد العينة أن الوسط الحسابي لعدد الكلمات الصحيحة يساوي ٣١ كلمة في الدقيقة الواحدة. ولقد دلت التجربة على أن الوسط لخريجي هذا البرنامج هو ٢٩ كلمة بانحراف معياري يساوي ٤. فهل هناك دليل كافٍ بدرجة ثقة ٩٥٪، على أن هناك تحسناً في مستوى خريجي هذا البرنامج؟

الحل:

ف : و = ٢٩

ف$_{١}$: و > ٢٩

أ = ٥٪

∴ ي$_{١-أ}$ = ي$_{٠,٩٥}$ = ١,٦٤ من جدول التوزيع الطبيعي بالملحق.

و = ٢٩

م = ٤

ن = ٢٥

سَ = ٣١

إذاً إحصائية الاختبار

$$ي = \frac{سَ - و}{م / \sqrt{ن}}$$

$$= \frac{٣١ - ٢٩}{٤ / \sqrt{٢٥}}$$

$$= ٢,٥$$

وبما أن القيمة المحسوبة (٢,٥) أكبر من المجدولة (١,٦٤) فالفرق بين الوسطين فرق جوهري (معنوياً) بدرجة ثقة ٩٥٪. لذلك لا يمكن قبول فرضية العدم، بمعنى أن هذا دليل على أن هناك ارتفاعاً في مستوى خريجي هذا البرنامج.

مثال (٨,٦) :

هل يمكن اعتبار وسط العينة لا يختلف عن وسط المجتمع فى المثال السابق؟

الحل :

ف. : و = ٢٩

$ف_١$: و ≠ ٢٩

ى = ٢,٥

إلا أن الفرضية البديلة ذات اتجاهين ؛ إذاً

أ = ٥٪

$\frac{أ}{٢}$ = ٢,٥٪

$ى_{١-\frac{أ}{٢}}$ = $ى_{٠,٩٧٥}$

= ١,٩٦

وبما أن

$ى > ى_{١-\frac{أ}{٢}}$

فالفرق بين الوسطين فرق جوهرى أيضاً لا يمكن أن يعزى للصدفة (Chance) ؛ لذلك لا يمكن قبول فرضية العدم .

(ب) عندما يكون تباين المجتمع الذى سحبت منه العينة مجهولاً ولكن حجم العينة كبير :

يعتبر الحجم الكبير للعينة مبرراً لاستبدال التباين المجهول للمجتمع ($م^٢$) بتباين العينة ($ع^٢$) . وفيما عدا ذلك تظل بقية قيم الاختبار كما كانت عليه فى الحالة السابقة ، بمعنى أن :

$$ى = \frac{\bar{س} - و}{\frac{ع}{\sqrt{ن}}} \qquad (٢)$$

مثال (٧,٨) :

تم اختيار عينة عشوائية حجمها ٦٤ من خريجى برنامج للنسخ الإعدادى، فكان الوسط الحسابى لعدد الكلمات الصحيحة فى الدقيقة الواحدة ٣١ كلمة بتباين قدره ١٦ ، ولقد دلت التجارب السابقة على أن الوسط لخريجى مثل هذا البرنامج هو ٢٩ كلمة . فهل هناك زيادة جوهرية بمستوى معنوية ٥٪ فى عدد الكلمات الصحيحة بالنسبة لخريجى هذا البرنامج؟

الحل :

$ف_٠$: و = ٢٩

$ف_١$: و > ٢٩

أ = ٠,٠٥

١ - أ = ٠,٩٥

∴ القيمة الحرجة هى (من جدول التوزيع الطبيعى بالملحق) :

$ى_{٠,٩٥}$ = ١,٦٤

ن = ٦٤

$ع^٢$ = ١٦

$\bar{س}$ = ٣١

إحصائية الاختبار هى :

$$ى = \frac{\bar{س} - و}{\frac{ع}{\sqrt{ن}}} \qquad (٢)$$

$$= \frac{٣١ - ٢٩}{\frac{٤}{٨}}$$

$$= ٤$$

∴ ى > $ى_{٠,٩٥}$

فالفرق جوهرى، ولا يمكن قبول فرضية العدم القائلة بأنه لا يوجد فرق بين هذا البرنامج وبقية البرامج السابقة. وربما يلاحظ هنا أن إحصائية الاختبار تتزايد بتزايد حجم العينة إذا لم تتغير بقية الإحصائيات.

(جـ) إذا كان حجم العينة صغيراً وتباين المجتمع مجهولاً :

إذا كان المجتمع الذى سحبت منه العينة طبيعياً، وكان حجم العينة صغيراً بدرجة لا تبرر استبدال تباين المجتمع بتباينها، فتوزيع إحصائية الاختبار هو توزيع تاء على (ن ـ ١) درجات حرية، بدلاً من التوزيع الطبيعى حسب ما جاء فى الفصل الخاص بالتوزيعات. بمعنى أن إحصائية الاختبار هى :

$$ت = \frac{\bar{س} - و}{\frac{ع}{\sqrt{ن}}} \qquad (٣) \quad \text{وذلك على (ن ـ ١) درجات حرية.}$$

أما القيمة الحرجة (المجدولة) فتستخرج من جدول توزيع تاء بالملحق حسب مستوى المعنوية ودرجات الحرية.

مثال (٨،٨) :

اختيرت عينة قوامها ١٦ خريجاً من أحد برامج النسخ، وكان الوسط الحسابى لعدد الكلمات الصحيحة فى الدقيقة الواحدة ٣١ كلمة بانحراف معيارى ٦. هل يمكن اعتبار هذا الفرق جوهرياً بمستوى معنوية ٥٪ مقارنة بالوسط لخريجى هذا البرنامج الذى أثبتت التجربة أنه يساوى ٢٩؟

الحل :

$ف_٠ : و = ٢٩$

$ف_١ : و > ٢٩$

$أ = ٠,٠٥$

$١ - أ = ٠,٩٥$

القيمة الحرجة من جدول توزيع تاء على ١٥ درجة حرية هى :

$ت_{٠,٩٥} = ١,٧٤٦$

$ن = ١٦$

$\sqrt{ن} = ٤$

$ع = ٦$

$\bar{س} = ٣١$

إحصائية الاختبار هى :

$$ت = \frac{\bar{س} - و}{\frac{ع}{\sqrt{ن}}} \qquad (٣)$$

$$= \frac{٣١ - ٢٩}{\frac{٦}{٤}}$$

$$= ١,٣٣٣$$

إذاً فالفرق ظاهرى، وليس جوهرياً ؛ لأن القيمة المحسوبة أقل من المجدولة ، بمعنى أنه يجب قبول فرضية العدم بمستوى ثقة ٩٥٪.

٤ ـ اختبارات الفرق بين وسطين من عينتين مستقلتين :

هنـاك حالات كثيرة فى المجـالات التـطبيقية التى تستـدعى اختبار الفرق بين وسطى مجتمعـين، اعتـماداً على القيم المستخـرجة من عينتين مستقلتين، فإذا كانت $و_١$ تمثل وسط المجتمع الأول، وكانت $و_٢$ تمثل وسط المجتمع الثانى ، فهناك عدد من الفرضيات التى تكون واحدة منها مجالاً للاختبار، فقد تكون فرضية العدم هى :

$ف : و_١ = و_٢$

أو بمعنى آخر أن الوسطين متساويان ، أو الفرق بينهما ليس جوهرياً ، بمعنى أن :

$ف : و_١ - و_٢ = صفراً$

وتكون فرضية العدم هنا إما على النحو الآتى :

$ف_{٠} : و_{١} > و$

أو :

$ف_{١} : و_{١} < و_{٢}$

أو بمعنى آخر :

$ف_{١} - ف_{٢} > صفر$

أو :

$ف_{٢} - ف_{١} > صفر$

وفى جميع هذه الحالات تكون على طرف واحد، وتعنى أن هناك وسطاً أكبر من الآخر.

كذلك قد تكون الفرضية البديلة على النحو التالى :

$ف_{١} : و_{١} \neq و_{٢}$

وهى فرضية ذات طرفين تعنى أن هناك فرقاً جوهرياً بين الوسطين.

أما إذا كان الفرق المحدد بفرضية العدم غير الصفر - يساوى ل مثلاً حيث ل قيمة معلومة - بمعنى أن :

$ف_{٠} : و_{١} - و_{٢} = ل$ وهى فرضية عدم بسيطة.

أو كانت مزدوجة على النحو الآتى :

$ف_{٠} : و_{١} - و_{٢} > ل$

فقد تكون الفرضية البديلة على النحو الاتى :

$ف_{١} : و_{١} - و_{٢} > ل$

أو

$ف_{١} : و_{١} - و_{٢} < ل$

أو

$ف_{١} : و_{١} - و_{٢} \neq ل$

واعتماداً على توزيع الفرق بين وسطين، فإحصائية الاختبار هى :

$$\text{ى} = \frac{(\bar{\text{س}}_1 - \bar{\text{س}}_2) - \text{ل}}{\sqrt{\frac{\text{م}_1^2}{\text{ن}_1} + \frac{\text{م}_2^2}{\text{ن}_2}}} \qquad (٤)$$

ولا تستخدم المعادلة السابقة إلا إذا كان تباين المجتمع الذى سحبت منه العينة الأولى (م_1^2) معلوماً، وكذلك تباين المجتمع الثانى (م_2^2).

هذا، وتبقى القيمة الحرجة (المجدولة) على ما كانت عليه اعتماداً على مستوى المعنوية . أما إذا كان التباين مجهولاً ، فيجوز استبدال تباين كل مجتمع بتباين عينته ($\text{ع}_{1,2}$)، إذا كان حجم العينة كبيراً . بمعنى أن :

$$\text{ى} = \frac{(\bar{\text{س}} - \bar{\text{س}}_2) - \text{ل}}{\sqrt{\frac{\text{ع}_1^2}{\text{ن}_1} + \frac{\text{ع}_2^2}{\text{ن}_2}}} \qquad (٤)$$

مثال (٨,٩) :

أجرى بحث لتقويم طريقتين فى التدريب على النسخ، واختيرت عينتان عشوائيتان من المتدربين على كل طريقة. اتضح أن الوسط الحسابى لعدد الكلمات الصحيحة فى الدقيقة الواحدة يساوى ٣٠، و ٣٢ فى المجموعة الأولى والثانية على التوالى. فإذا كان حجم الأولى ٤٥ ، والثانية ٣٦ متدرباً، بينما كان التباين لعينة المتدربين بالطريقة الأولى يساوى ٨١، وللمتدربين بالطريقة الثانية يساوى ١٤٤ ، فهل هناك فرق جوهرى بمستوى معنوية ١٠٪ بين الطريقتين؟

الحل :

ف : $\text{و}_1 = \text{و}_2$ أى أن الفرق بين الطريقتين ليس جوهرياً. أو بمعنى آخر :

ف : $\text{و}_1 - \text{و}_2 = $ صفراً

$\text{ف}_{١}$: $\text{و}_{١} \neq \text{و}_{٢}$ وهى فرضية بديلة ذات طرفين.

$\text{أ} = ٠,١٠$

$\frac{\text{أ}}{٢} = ٠,٠٥$

$\text{ى}_{١ - \frac{\text{أ}}{٢}} = \text{ى}_{٠,٩٥}$

إذاً القيمة الحرجة من جدول التوزيع الطبيعى

$\text{ى}_{٠,٩٥} = ١,٦٤$

إحصائية الاختبار هى :

$$\text{ى} = \frac{(\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}) - \text{ل}}{\sqrt{\frac{\text{ع}_{١}^{٢}}{\text{ن}_{١}} + \frac{\text{ع}_{٢}^{٢}}{\text{ن}_{٢}}}} \qquad (٤)$$

وذلك لأن حجم العينة كبير.

$\bar{\text{س}}_{١} = ٣٠$

$\bar{\text{س}}_{٢} = ٣٢$

ل = صفراً من فرضية العدم.

$\text{ن}_{١} = ٤٥$

$\text{ن}_{٢} = ٣٦$

$\text{ع}_{١}^{٢} = ٨١$

$\text{ع}_{٢}^{٢} = ١٤٤$

$$\therefore \text{ى} = \frac{٣٠ - ٣٢}{\sqrt{\frac{٨١}{٤٥} + \frac{١٤٤}{٣٦}}}$$

$$= -٠,٨٣$$

وبما أن $-٠,٨٣ > -١,٦٤$

فإن الفرق بين الطريقتين ظاهرى وليس جوهرياً، بمعنى أنه لا يمكن رفض فرضية العدم القائلة بأنه لا يوجد فرق بين الطريقتين.

البرنامج التالى يحسب إحصائية الاختبار للفرق بين وسطين حسب ما هو وارد بالمثال (٨,٩) حيث:

N1 حجم العينة الأولى.
N2 حجم العينة الثانية.
X1 الوسط الحسابى للعينة الأولى.
X2 الوسط الحسابى للعينة الثانية.
V1 التباين للعينة الأولى.
V2 التباين للعينة الثانية.
L الفرق بين الوسطين، وفى هذه الحالة فهو صفر نسبة لافتراض أنه لا يوجد فرق.

```
10 REM                  برنامج لاختبار الفرق بين وسطين
20 READ N1,N2,X1,X2,V1,V2,L
30 DATA 45,36,30,32,81,144,0
40 Y=(X1-X2-L)/SQR(V1/N1+V2/N2)
50 PRINT TAB(20);Y;' = احصائيه الاختبار '
60 END

                    المخرجات
                    ---------

احصائيه الاختبار  =     -.8304549
```

وأما إذا كان حجم العينة صغيراً، والتباين مجهولاً، بافتراض أن تبايني المجتمعين اللذين سحبت منها العينتان متساويان، فإحصائية الاختبار - راجع الفصل الخاص بأهم التوزيعات - هى:

$$\text{ت}_{(\text{ن}_1+\text{ن}_2-2)} = \frac{(\bar{\text{س}}_1 - \bar{\text{س}}_2) - \text{ل}}{\sqrt{\frac{(\text{ن}_1-1)\text{ع}_1^2 + (\text{ن}_2-1)\text{ع}_2^2}{\text{ن}_1+\text{ن}_2-2}}\ \sqrt{\frac{1}{\text{ن}_1}+\frac{1}{\text{ن}_2}}} \qquad (٥)$$

مثال (٨,١٠):

أجرى بحث لتقويم طريقتين فى التدريب على النسخ، ولقد تم اختيار عينتين عشوائيتين من مجموعتين تدربت كل منها على واحدة من الطريقتين. كان حجم العينة الأولى ٥ أفراد،

والثانية ٦ أفراد، بينما كان الوسط الحسابى لعدد الكلمات الصحيحة فى الدقيقة الواحدة يساوى ٣٠ للمجموعة الأولى، و ٣٢ للمجموعة الثانية . أما التباين الخاص بالقيم العينية للمجموعتين فيساوى ٩ و ١١ للمجموعة الأولى والثانية على التوالى .

فهل هناك فرق جوهرى بمستوى معنوية ١٠٪ بين الطريقتين؟

الحل :

$ف_{٠} : و_{١} = و_{٢}$

$ف_{١} : و_{١} \neq و_{٢}$

$أ = ١٠\%$

$\frac{أ}{٢} = ٥\%$

$ن_{١} = ٥$

$ن_{٢} = ٦$

∴ درجات الحرية $= ٥ + ٦ - ٢$

$= ٩$

القيمة الحرجة من جدول توزيع تاء بالملحق وبمستوى ثقة ٠,٩٥ وعلى ٩ درجات حرية هى :

$$ت_{(٩،٠,٩٥)} = ١,٨٣٣$$

إحصائية الاختبار (المحسوبة) هى :

$$ت = \frac{(\bar{س}_{١} - \bar{س}_{٢}) - ل}{\sqrt{\frac{(ن_{١} - ١)ع_{١}^{٢} + (ن_{٢} - ١)ع_{٢}^{٢}}{ن_{١} + ن_{٢} - ٢}}\sqrt{\frac{١}{ن_{١}} + \frac{١}{ن_{٢}}}} \qquad (٥)$$

$$= \frac{(٣٠ - ٣٢) - ٠}{\sqrt{\frac{٤ \times ٩ + ٥ \times ١١}{٥ + ٦ - ٢}}\sqrt{\frac{١}{٥} + \frac{١}{٦}}}$$

$$= -١,٠٣٩$$

وبما أن $-١,٠٣٩ > -١,٨٣٣$

فالفرق ظاهرى وليس جوهرياً، وعليه لا بد من قبول فرضية العدم القائلة بأنه لا يوجد فرق بين الطريقتين.

البرنامج التالى يقوم بحساب إحصائية الاختبار حسب ما هو وارد بالمثال (١٠ ,٨) السابق باستخدام المعادلة

$$T = \frac{A}{\sqrt{\frac{B}{C} \times \frac{1}{D}}}$$

حيث :

$A = X_1 - X_2 - L$ = الفرق

$L = 0$

$B = (N_1 - 1)V_1 + (N_2 - 1)V_2$

$C = N_1 + N_2 - 2$

$D = \frac{1}{N_1} + \frac{1}{N_2}$

N_1 = حجم العينة الأولى

N_2 = حجم العينة الثانية

V_1 = تباين العينة الأولى

V_2 = تباين العينة الثانية

```
10 REM        برنامج لاختبار الفرق بين وسطين
20 READ N1,N2,X1,X2,V1,V2,L
30 DATA 5,6,30,32,9,11,0
40 A=X1-X2-L
50 B=(N1-1)*V1+(N2-1)*V2
60 C=N1+N2-2
70 D=1/N1+1/N2
80 T=A/(SQR(B/C)*SQR(D))
90 PRINT TAB(20);T;" = احصائيه الاختبار "
100 END

                         المخرجات
                         --------

              -1.038712     = احصائيه الاختبار
```

٥ - اختبار الفرق بين وسطين لأزواج متشابهة أو لعينة واحدة :

تعتمـد جميع الاختبارات السابقة على عينتين مستقلتين، أو معلم واحد لعينة، ويحدث أحياناً أن يكون الهدف هو اختبار الفرق بين وسطين لتجربتين أجريتا على نفس المجموعة من أفراد العينة، أو لقياس الفرق بين الوسطين قبل وبعد إجراء تجربة معينة، أو قياس أثر التجربة بعد تقسيم أفراد العينة إلى عينتين بأزواج متشابهة. ومثال ذلك الفرق بين وسطى الدرجات لمادتين بالنسبة لنفس مجموعة الطلاب، أو الفرق بين الوسطين للسرعة قبل وبعد التدريب على النسخ لنفس المتدربين، أو الفرق بين وسطى الزمن الذى يستغرقه عقاران لتجلط الدم بعد تقسيم أفراد العينة إلى عينتين متشابهتين.

فإذا كانت (أ) هى التجربة الأولى، و (ب) هى التجربة الثانية، بينما $و_{(أ)}$ ، و $_{(ب)}$ هما وسطا المجتمعين، وبافتراض أن $(و_ر)$ هى قيمة معينة و (ل) هى الفرق، فالفرق بين أى زوجين متشابهين هو :

$$ل_ر = س_{أر} - س_{بر} \qquad (٦)$$

أما الوسط الحسابى للفروق $(\bar{ل})$ فهو :

$$\bar{ل} = \frac{١}{ن} \sum ل_ر \qquad (٧)$$

حيث (ن) هى عدد الأزواج.

أما تباين الفروق فهو بالتالى :

$$ع_ل^٢ = \frac{\sum (ل_ر - \bar{ل})^٢}{ن - ١} \qquad (٨)$$

وعليه تكون إحصائية الاختبار تابعة لتوزيع تاء على (ن - ١) درجات حرية، أى أنها :

$$ت_{(ن-١)} = \frac{(\bar{ل}) - (و_أ - و_ب)}{ع_ل / \sqrt{ن}} \qquad (٩)$$

وباعتبار أن فرضية العدم هى :

$$ف_ب : و_أ = و_ب$$

فإحصائية الاختبار هى :

$$ت_{(ن - ١)} = \frac{\bar{ل}}{ع_ل / \sqrt{ن}} \qquad (١٠)$$

مثال (١١, ٨) :
أجريت دراسة لقياس الأثر لطريقة معينة فى التدريب على النسخ، فاختيرت عينة من ٢٥ متدرباً، وتم رصد السرعة فى الدقيقة الواحدة لكل متدرب قبل التحاقه بالبرنامج وبعد انتهاء فترة التدريب على تلك الطريقة، فكان مجموع الفروقات بين السرعتين يساوى ١٢٥، بينما كان مجموع مربعات تلك الفروقات ٨٥٠.

فهل تعتبر هذه الطريقة مفيدة بمستوى معنوية ١٠٪؟

الحل :

$ف_٠ : و_أ = و_ب$

$ف_١ : و_أ \neq و_ب$ لأن الفرق غير محدد.

$أ = ١٠\%$

$\frac{أ}{٢} = ٥\%$

القيمة الحرجة ـ باستخدام جدول توزيع تاء على ٢٤ درجة حرية وبمستوى ٩٥٪ هى :

$$ت_{(٢٤، ٠،٩٥)} = ١,٧١١$$

إحصائية الاختبار هى :

$$ت = \frac{\bar{ل}}{ع_ل / \sqrt{ن}} \qquad (١٠)$$

$$\bar{ل} = \frac{\sum ل_ر}{ن} \qquad (٧)$$

$$ع_ل^٢ = \frac{\sum ل_ر^٢ - \frac{(\sum ل_ر)^٢}{ن}}{ن - ١} \qquad (٨)$$

$$\sum \text{ل}_{\text{ر}} = ١٢٥$$

$$\sum \text{ل}^{٢}_{\text{ر}} = ٨٥٠$$

$$\therefore \bar{\text{ل}} = \frac{١٢٥}{٢٥}$$

$$= ٥$$

$$\text{ع}^{٢}_{\text{ر}} = \frac{٨٥٠ - \frac{١٢٥ \times ١٢٥}{٢٥}}{٢٤}$$

$$= ٩,٣٧٥$$

$$\therefore \text{ع}_{\text{ر}} = ٣,٠٦٢$$

$$\text{ت} = \frac{٥}{٣,٠٦٢ / \sqrt{٢٥}}$$

$$= ٨,١٦٥$$

وبما أن

$$٨,١٦٥ > ١,٧١١$$

فلا يمكن قبول فرضية العدم، بمعنى أنه يجب قبول الفرضية البديلة القائلة بأن هناك فرقاً بين السرعتين، ولا يمكن إنكار أثر تلك الطريقة.

فيما يلى برنامج لحل المسألة فى المثال (٨،١١) باستخدام المعادلتين :

$$M = (D_2 - (D_1 \times D_1 / N) / (N - 1))$$

$$T = \frac{D_1}{N \sqrt{M/N}}$$

حيث :

T = إحصائية الاختبار

N = حجم العينة

D_1 = مجموع الفروقات

D_2 = مجموع مربعات الفروقات

M = $\text{ع}^{٢}_{\text{ل}}$

```
10 REM      اختبار الفرضيات
20 READ N,D1,D2
30 DATA 25,125,850
40 M=(D2-(D1*D1/N))/(N-1)
50 T=D1/N/(SQR(M)/SQR(N))
60 PRINT TAB(20);T;" = احصائيه الاختبار '
70 END
                    المخرجات
                    --------
        8.164966      = احصائيه الاختبار
```

٦ ـ اختبار الفرق لأكثر من وسطين :

يتم اختبار الفرق لأكثر من وسطين بأسلوب تحليل التباين الذى سيأتى ذكره كاملاً للحالات غير المعلمية بالفصل القادم.

٧ ـ اختبارات النسب :

أ ـ اختبار نسبة واحدة :

اتضح من الفصل الخاص بأهم التوزيعات أن النسبة الخاصة بصفة معينة ـ مثل نسبة الذين أجابوا بنعم على سؤال معين ـ تتبع التوزيع ذا الحدين، فإذا كانت حَ تعنى نسبة أفراد العينة الذين يتمتعون بصفة معينة، وكانت ح هى نسبة أفراد المجتمع الذين يتمتعون بتلك الصفة، فإحصائية الاختبار تتبع التوزيع الطبيعى ـ راجع الفصل الخاص بأهم التوزيعات ـ على النحو التالى :

$$ى = \frac{حَ - ح}{\sqrt{ح(1-ح)/ن}} \qquad (11)$$

مثال (٨,١٢) :

يعتقد أن مقدرة الرجال فى حل المسائل الرياضية أكبر من مقدرة النساء، فاختيرت عينة عشوائية من ٣٠٠ شخص من بين الناجحين فى امتحان موحد للرياضيات، فاتضح أن بينهم ١٦٢ رجلاً والباقى من النساء. فهل يعتبر ذلك دليلاً كافياً بدرجة ثقة ٩٥٪ لاعتبار أن الرجال أكبر مقدرة من النساء فى حل المسائل الرياضية؟

الحل :

لو أن عدد الـرجال مساوٍ لعدد النساء، لأصبحت نسبة الرجال مساوية لنسبة النساء، ويساوى كل منهما ٠,٥ ، لذلك فإن :

$$\text{ف}_{\cdot} : \text{ح} = ٠,٥$$

$$\text{ف}_{١} : \text{ح} > ٠,٥$$

$$١ - \text{أ} = ٠,٩٥$$

∴ القيمة الحرجة من جدول التوزيع الطبيعى هى :

$$\text{ى}_{٠,٩٥} = ١,٦٥$$

إحصائية الاختبار هى :

$$\text{ى} = \frac{\hat{\text{ح}} - \text{ح}}{\sqrt{\frac{\text{ح}(١ - \text{ح})}{\text{ن}}}} \qquad (١١)$$

$$= \frac{\frac{١٦٢}{٣٠٠} - ٠,٥}{\sqrt{\frac{٠,٥ \times ٠,٥}{٣٠٠}}}$$

$$= ١,٣٨٦$$

وبما أن :

$$١,٦٥ > ١,٣٨٦$$

فلا بد من قبول فرضية العدم، وهذا يعنى أن ذلك ليس كافياً لاعتبار الرجال أكثر كفاءة من النساء فى حل المسائل الرياضية.

البرنامج التالى يحسب إحصائية الاختبار للمسألة بالمثال (١٢). المتغيرات المستخدمة :

N حجم العينة
N_1 تكرار الصفة المطلوب اختبارها
C النسبة حسب فرضية العدم

أما إحصائية الاختبار فهى :

$$M = \frac{\left(\frac{N_1}{N} - C\right)}{\sqrt{\frac{C(1-C)}{N}}}$$

```
10 REM        اختبار الفرضيات
20 READ N,N1,C
30 DATA 300,162,0.5
40 M=(N1/N-C)/(SQR(C*(1-C)/N))
50 PRINT TAB(20);M;' = احصائيه الاختبار'
60 END
                              المخرجات
                              ---------

               1.385639     = احصائيه الاختبار
```

(ب) اختبار الفرق بين نسبتين لعينتين كبيرتين :

إذا كانت حَ١ وحَ٢ تمثلان نسبتين من عينتين مستقلتين حجماهما ن١ ون٢ على التوالى، وإذا كان حجما العينتين كبيرين، بينما كانت ح١ هى نسبة المجتمع الأول الذى يتمتع بصفة معينة، تقابلها حَ٢ من المجتمع الثانى، وبافتراض أن لَ هى الفرق بين النسبتين بناء على الفرضية

ف : ح١-ح٢ = لَ

فإحصائية الاختبار تتبع التوزيع الطبيعى على النحو الآتى :

$$ى = \frac{(\acute{ح}_1 - \acute{ح}_2) - ل}{\sqrt{\frac{\bar{ح}(1-\bar{ح})}{ن_1} + \frac{\bar{ح}(1-\bar{ح})}{ن_2}}} \quad (12)$$

حيث :

$$\bar{ح} = \frac{ن_1 \acute{ح}_1 + ن_2 \acute{ح}_2}{ن_1 + ن_2} \quad (13)$$

مثال (١٣,٨) :

مصنعان لإنتاج المصابيح الكهربائية يعتقد أنهما متساويان من حيث الصفات فى الإنتاج. اختيرت عينة من ٢٠٠٠ مصباح من المصنع الأول، و ٣٠٠٠ من المصنع الثانى، فكان عدد المصابيح التى رفضت بعد فحصها يساوى ٥٠٠ مصباح من إنتاج المصنع الأول، و ٩٠٠ من الثانى، فهل يعتبر ذلك دليلاً كافياً بمستوى معنوية ٥٪ على أن إنتاج المصنع الأول أفضل من الثانى؟

الحل :

$ف : ح_1 = ح_2$

$ف_1 : ح_1 < ح_2$

حيث ح هى نسبة التالف.

أ = ٥٪

١ ـ أ = ٠,٩٥

∴ القيمة الحرجة (المجدولة) هى :

$ى_{0.95} = 1.65$

إحصائية الاختبار هى :

$$\text{ى} = \frac{(\hat{\text{ح}}_{١} - \hat{\text{ح}}_{٢})}{\sqrt{\frac{\hat{\text{ح}}(١ - \hat{\text{ح}})}{\text{ن}_{١}} + \frac{\hat{\text{ح}}(١ - \hat{\text{ح}})}{\text{ن}_{٢}}}} \qquad (١٢)$$

$$\hat{\text{ح}} = \frac{\text{ن}_{١}\hat{\text{ح}}_{١} + \text{ن}_{٢}\hat{\text{ح}}_{٢}}{\text{ن}_{١} + \text{ن}_{٢}} \qquad (١٣)$$

$$\hat{\text{ح}}_{١} = \frac{٥٠٠}{٢٠٠٠}$$

$$= ٠,٢٥$$

$$\hat{\text{ح}}_{٢} = \frac{٩٠٠}{٣٠٠٠}$$

$$= ٠,٣٠$$

$$\therefore \hat{\text{ح}} = \frac{٩٠٠ + ٥٠٠}{٣٠٠٠ + ٢٠٠٠}$$

$$= ٠,٢٨$$

$$١ - \hat{\text{ح}} = ٠,٧٢$$

$$\text{ى} = \frac{٠,٢٥ - ٠,٣٠}{\sqrt{\frac{٠,٢٨ \times ٠,٧٢}{٢٠٠٠} + \frac{٠,٢٨ \times ٠,٧٢}{٣٠٠٠}}}$$

$$\therefore \text{ى} = -٣,٨٤٦$$

وبما أن $-٣,٨٤٦ < -١,٦٥$

فلا بد من رفض فرضية العدم وقبول الفرضية البديلة ؛ لأن إحصائية الاختبار تقع فى منطقة الرفض. وهذا يعنى أنه وبمستوى معنوية ٥٪ يمكن القول بأن إنتاج المصنع الأول أفضل من الثانى من حيث المواصفات.

فيما يلى برنـامـج لحسـاب إحصائية الاختبار للمسألة الواردة فى المثال (٨,١٣) السابق الخاصة باختبار الفرق بين نسبتين لعينتين كبيرتين باستخدام المعادلة :

$$M = \frac{L_1 - L_2}{\sqrt{L(1-L)/N_1) + L(1-L)/N_2}}$$

$$L = \frac{X_1 - X_2}{N_1 + N_2}$$

X_1 = تكرار الظاهرة الأولى

X_2 = تكرار الظاهرة الثانية

N_1 = المجتمع الأول

N_2 = المجتمع الثانى

$L_1 = X_1/N_1$

$L_2 = X_2/N_2$

M = إحصائية الاختبار

```
10 REM        اختبار الفرضيات
20 READ N1,N2,X1,X2
30 DATA 2000,3000,500,900
40 L1=X1/N1
50 L2=X2/N2
60 L=(X1+X2)/(N1+N2)
70 M=(L1-L2)/SQR(L*(1-L)/N1+L*(1-L)/N2)
80 PRINT TAB(20);M;' = إحصائية الاختبار'
90 END

              -3.857581     = إحصائية الاختبار
```

(جـ) اختبار الفرق بين نسبتين لعينتين صغيرتين :

اتضح مما مضى أنـه إذا تم اختيار عينتين عشوائيتين مستقلتين من مجتمعين طبيعيين، فتوزيع الفرق بين وسطيهما يتبع توزيع تاء. ولقد لوحظ أيضاً التشابه بين توزيعات النسب والأوساط، فتوزيع النسبة مشابه لتوزيع الوسط، وتوزيع الفرق بين نسبتين يشابه توزيع الفرق بين وسطين.

فإذا كانت فرضية العدم على النحو التالى :

$$\text{ف}_0 : \text{ح}_1 - \text{ح}_2 = \bar{\text{ل}}$$

فإحصائية الاختبار هى :

$$\text{ت} = \frac{(\hat{\text{ح}}_1 - \hat{\text{ح}}_2) - \bar{\text{ل}}}{\sqrt{\frac{\hat{\text{ح}}_1 (1 - \hat{\text{ح}}_1)}{\text{ن}_1} + \frac{\hat{\text{ح}}_2 (1 - \text{ح}_2)}{\text{ن}_2}}} \qquad (١٤)$$

على (ن_1 ـ ١) أو (ن_2 ـ ١) درجات حرية أيهما أكبر.

مثال (٨,١٤) :

يعتقد أن الرجال أكثر قدرة من النساء على حل المسائل الرياضية، فاختيرت عينة من ١٥ رجـلاً و ٢٠ امرأة بنفس المستوى التعليمى، وجلسوا لامتحان موحد فى الجبر. اتضح بعد النتيجة أن عدد الراسبين يساوى ثمانية بينهم ٣ رجال فقط. فهل يعتبر هذا دليلاً كافياً بمستوى معنوية ٢,٥٪ على أن مستوى الرجال فى الرياضيات أفضل من النساء؟

الحل :

$\text{ف}_0 : \text{ح}_1 = \text{ح}_2$ نسب الراسبين متساوية

$\text{ف}_1 : \text{ح}_1 < \text{ح}_2$ نسبة الراسبين بين الرجال أقل من النساء

$\text{ن}_1 - ١ = ١٤$

$\text{ن}_2 - ١ = ١٩$

$$أ = ٢,٥\%$$

$$\therefore ١ - أ = ٠,٩٧٥$$

∴القيمة الحرجة (المجدولة) من توزيع تاء على ١٩ درجة حرية وبدرجة ثقة ٠,٩٧٥ هى :

$$ت = ٢,٠٩٣$$

إحصائية الاختبار هى :

$$ت = \frac{ح'_{١} - ح'_{٢}}{\sqrt{\frac{ح'_{١}(١-ح'_{١})}{ن_{١}} + \frac{ح'_{٢}(١-ح'_{٢})}{ن_{٢}}}} \qquad (١٤)$$

حيث :

$$ح'_{١} = \frac{٣}{١٥}$$

$$= ٠,٢٠$$

$$١ - ح'_{١} = ٠,٨٠$$

$$ح'_{٢} = \frac{٥}{٢٠}$$

$$= ٠,٢٥$$

$$١ - ح'_{٢} = ٠,٧٥$$

$$ت = \frac{٠,٢٠ - ٠,٢٥}{\sqrt{\frac{٠,٢٠ \times ٠,٨٠}{١٥} + \frac{٠,٢٥ \times ٠,٧٥}{٢٠}}}$$

$$= - ٠,٣٥٢$$

وبما أن ـ ٠,٣٥٢ > ـ ٢,٠٩٣ فإحصائية الاختبار ضمن منطقة القبول. إذاً لا بد من قبول فرضية العدم، وعليه ليس هناك دليل بمستوى ٢,٥٪ على تفوق الرجال على النساء فى حل المسائل الرياضية.

فيما يلى برنامج لإيجاد إحصائية الاختبار للبيانات الواردة فى المثال (٨ , ١٤) السابق الخاصة باختبار الفرضية لعينتين صغيرتين باستخدام إحصائية الاختبار

$$M = \frac{B_1 - B_2}{D}$$

حيث :

$M =$ ت

$B_1 =$ ح/١ = النسبة الأولى

$B_2 =$ ح/٢ = النسبة الثانية

$$D = \sqrt{C_1 + C_2}$$

$C_1 = B_1 (1 - B_1)/N_1$

$C_2 = B_2 (1 - B_2)/N_1$

```
10 REM        اختبار الفرضيات
20 READ N1,N2,A1,A2
30 DATA 15,20,3,5
40 B1=A1/N1
50 B2=A2/N2
60 C1=B1*(1-B1)/N1
70 C2=B2*(1-B2)/N2
80 D=SQR(C1+C2)
90 M=(B1-B2)/D
100 PRINT TAB(20);M;' = احصائيه الاختبار'
110 END
                         المخرجات
                         --------
              -.3531859     = احصائيه الاختبار
```

٨ - اختبارات التباين :

أ - اختبار التباين لعينة واحدة :

يكـون الهـدف من الاختبـار فى بعض الحـالات هو اتخاذ قرار بشأن التجانس بين قيم المجتمع . فزيادة التباين فى أى صفة من الصفات الخاصة بإنتاج أحد المصانع مثلاً، تدل على

ضعف استمرارية الدقة، فإذا كان تباين المجتمع هو م^{2}، يجب ألا يكون الفرق بينه وبين المقدار المحدد للتباين (م_{0}^{2}) فرقاً جوهرياً، بمعنى أن فرضية العدم هى :

$$\text{ف}_{0} : \text{م}^{2} = \text{م}_{0}^{2}$$

وبما أن التباين الذى يمكن قياسه بالفعل، واتخاذه تقديراً لتباين المجتمع، هو تباين العينة (ع^{2})، فلا مناص من الاعتماد عليه فى اتخاذ القرار. ولقد اتضح من الفصل الخاص بأهم التوزيعات أن :

[$(\text{ن} - 1)\,\text{ع}^{2} / \text{م}_{0}^{2}$] تتبع توزيع مربع كاى على (ن - ١) درجات حرية . وعليه تكون إحصائية الاختبار هى :

$$\text{ك} = \frac{(\text{ن} - 1)\,\text{ع}^{2}}{\text{م}_{0}^{2}} \qquad (15)$$

بينما تكون القيمة الحرجة هى القيمة المستخرجة من توزيع مربع كاى على (ن - ١) درجات حرية. بيد أن القيم المكونة لإحصائية الاختبار لا تكون إلا موجبة فى جميع الحالات، كما أن توزيع مربع كاى غير متشابه؛ لذلك يكون الحد الأعلى لمنطقة القبول عند مستوى المعنوية (أ)، إذا كانت الفرضية البديلة ذات طرف واحد والمستوى ($\frac{\text{أ}}{2}$) إذا كانت ذات طرفين. أما الحد الأدنى لمنطقة القبول فتكون عند مستوى معنوية (١ - أ)، إذا كانت الفرضية البديلة ذات طرف واحد، وبالمستوى ($1 - \frac{\text{أ}}{2}$) إذا كانت الفرضية البديلة ذات طرفين.

مثال (٨,١٥) :

يتم التقدير النهائى للمتدربين بمعهد الإدارة العامة بافتراض أن الوسط الحسابى لدرجات الممتحنين فى كل مادة يساوى ٧٥، والانحراف المعيارى يساوى ١٠ ، اختيرت عينة عشوائية من ٢٥ متدرباً فى أحد البرامج فاتضح أن الوسط الحسابى لدرجات أفراد هذه العينة فى مادة مبادىء الإدارة يساوى بالفعل ٧٥، إلا أنه اتضح أن الانحراف المعيارى يساوى ١٢ ، فهل يعتبر الانحراف المعيارى فى هذه المادة أكبر من الانحراف المعيارى المحدد لدرجة تؤثر على دقة التقويم، علماً بأن مستوى معنوية الاختبار يساوى ٥٪؟

الحل :

$ف_{\cdot}$: $م^٢$ = $م_{\cdot}^٢$

أى أن :

$ف_{\cdot}$: $م^٢$ = ١٠٠

$ف_{١}$: $م^٢$ > ١٠٠

أ = ٥٪

١ ـ أ = ٩٥٪

ن ـ ١ = ٢٤

∴ الحد الأعلى لمنطقة القبول (القيمة الحرجة) من جدول توزيع مربع كاى على ٢٤ درجة حرية يكون بمستوى ٠,٠٥ وهو يساوى

ك = ٣٦,٤١٥

(لاحظ أن الحد الأدنى يكون بمستوى ٠,٩٥ ويساوى ١٣,٨٤٨٤).

أما إحصائية الاختبار فهى :

$$ك = \frac{(ن - ١)ع^٢}{م_{\cdot}^٢} \qquad (١٥)$$

$$= \frac{٢٤ \times (١٢)^٢}{(١٠)^٢}$$

$$= ٣٤,٥٦$$

وبما أن إحصائية الاختبار أقل من الحد الأعلى لمنطقة القبول، فيجب قبول فرضية العدم بمستوى معنوية ٥٪. إذاً فلا يوجد فرق جوهرى بين التباين فى تلك المادة والتباين المحدد للتقويم.

البرنامج التالى يقوم بحساب إحصائية الاختبار للتباين بين صفتين للبيانات الواردة فى المثال (٨,١٥) السابق.

المتغيرات المستخدمة :

N	حجم العينة
X_1	الوسط الحسابى للصفة الأولى
V_1	الانحراف المعيارى للصفة الأولى
X_2	الوسط الحسابى للصفة الثانية
V_2	الانحراف المعيارى للصفة الثانية

```
10 REM        اختبار الفرضيات
20 READ N,X1,V1,X2,V2
30 DATA 25,75,10,75,12
40 M=((N-1)*V2*V2)/V1**2
50 PRINT TAB(20);M;' = احصائيه الاختبار'
60 END
                              المخرجات
                              --------

                     34.56   = احصائيه الاختبار
```

(ب) اختبار المقارنة بين تباينين :

إذا كان الهدف من الاختبار هو مقارنة تباينين لمجتمعين طبيعيين مستقلين بعضهما عن بعض لمعرفة ما إذا كان الفرق بينهما جوهرياً أم ظاهرياً، أو لتحديد أيهما أكثر تجانساً باستمرار، ففرضية العدم هى :

$$ف : م_1^2 = م_2^2$$

بينما تتخذ الفرضية البديلة أحد الأنماط التالية :

$$ف_1 : م_1^2 > م_2^2$$

أو

$$ف_1 : م_1^2 < م_2^2$$

أو :

$$ف_١ : م_١^٢ \neq م_٢^٢$$

فإذا كانت فرضية العدم مقبولة فإن نسبة أي من التباينين تساوى الواحد أو قريبة منه ($ف : م_٢^٢ / م_١^٢ = ١$). أما إذا كان الفرق بين التباينين جوهرياً، فتكون النسبة بعيدة عن الواحد، ولكنها لا تكون سالبة؛ لأن التباين موجب فى جميع الحالات.

تعتمد إحصائية الاختبار على تقديرى التباينين وهما $ع_١^٢$ و $ع_٢^٢$، ولقد اتضح من الفصل الخاص بأهم التوزيعات أن :

$\frac{ع_١^٢}{ع_٢^٢}$ تتبع توزيع فاء على ($ن_١ - ١$) و ($ن_٢ - ١$) درجات حرية مما جعل البعض يسمى توزيع فاء بتوزيع نسبة التباين. وخلاصة القول هى أن إحصائية الاختبار :

$$ف = \frac{ع_١^٢}{ع_٢^٢} \qquad (١٦)$$

وأما القيمة الحرجة فتستخرج من جدول توزيع فاء بالملحق على ($ن_١ - ١$) و ($ن_٢ - ١$) - أى درجات حرية البسط أولاً، ثم درجات حرية المقام - وبمستوى المعنوية المحدد.

وتجدر الإشارة هنا إلى أن التباين الأكبر يكون بسطا إذا كانت الفرضية البديلة ذات طرف واحد، وذلك تفادياً للبحث عن أدنى حد للقبول. أما إذا كانت الفرضية البديلة ذات طرفين، فالحد الأدنى للقيمة الحرجة هو مقلوب الحد الأعلى للقيمة الحرجة بعد استبدال درجات حرية البسط والمقام؛ لأن القيم المجدولة هى القيم العليا فقط. إذاً فلاستخراج الحد الأدنى للقبول تستخرج القيمة المناظرة من جدول توزيع فاء على ($ن_٢ - ١$) و ($ن_١ - ١$) درجات حرية ومن ثم مقلوبها.

بيد أن توزيع فاء يتأثر كثيراً بالانحراف عن التوزيع الطبيعى؛ لذلك يجب عدم استخدامه فى حالة العينات الصغيرة الحجم، إلا إذا كان المجتمع الخاص بكل عينة طبيعياً.

مثال (٨,١٦) :

كان الانحراف المعيارى لدرجات ١٥ دارساً فى مادة الرياضيات يساوى ٢٠، بينما كان الانحراف المعيارى لدرجات ١٧ طالباً فى الاقتصاد يساوى ١٦ . فهل يعتبر ذلك دليلاً كافياً بمستوى معنوية ١٪ على أن مادة الرياضيات أكبر قدرة على التمييز بين الطلاب؟

الحل :

$$ف_{\cdot} : م_{١}^{٢} = م_{٢}^{٢}$$

$$ف_{١} : م_{١}^{٢} > م_{٢}^{٢}$$

$$أ = ٠,٠١$$

$$ن_{١} - ١ = ١٤$$

$$ن_{٢} - ١ = ١٦$$

القيمة الحرجة من جدول توزيع فاء هى :

$ف_{(١٤،١٦)}$ بمستوى ١٪ = ٣,٤٥

إحصائية الاختبار :

$$ف = \frac{ع_{١}^{٢}}{ع_{٢}^{٢}} \qquad (١٦)$$

$$\therefore ف = \frac{٢٠^{٢}}{١٦^{٢}}$$

$$= ١,٥٦٣$$

وبما أن إحصائية الاختبار أقل من القيمة الحرجة، فهى إذاً فى منطقة القبول. لذلك لا بد من قبول فرضية العدم القائلة بأنه لا يوجد فرق جوهرى بين مقدرة الرياضيات والاقتصاد فى التمييز بين الطلاب.

ملحوظة :

لو كانت الفرضية البديلة ذات طرفين ولو كان مستوى المعنوية ٢٪، لكان الحد الأعلى لمنطقة القبول هو :

$ف_{(١٦،١٠)}$ بمستوى معنوية ١٪ أيضاً = ٣,٤٥

ولإيجاد الحد الأدنى تستخرج أولاً

$ف_{(١٦،١٤)}$ بمستوى معنوية ١٪ = ٣,٦٢

ومقلوبها :

$$\frac{١}{ف_{(٠,١،١٦،١٤)}} = \frac{١}{٣,٦٢}$$

= ٠,٢٧٦

فيما يلي برنامج لحساب إحصائية اختبار المقارنة بين تباينين حسب البيانات الواردة بالمثال (٨,١٦) السابق وباستخدام المعادلة

$$M = \frac{(V_1)^2}{(V_2)^2}$$

حيث :

M = ف = إحصائية الاختبار

V_1 = الانحراف المعيارى للعينة الأولى

V_2 = الانحراف المعيارى للعينة الثانية

```
10 REM      اختبار الفرضيات
20 READ V1,V2
30 DATA 20,16
40 M=(V1*V1)/(V2*V2)
50 PRINT TAB(20);M; ' = احصائيه الاختبار '
60 END

                              المخرجات
                              --------
            1.5625   = احصائيه الاختبار
```

تمارين

١ ـ وضح بيانياً الفرق بين منطقة القبول ومنطقة الرفض.

٢ ـ ما هى أنواع القرارات التى تتخذ تحت ظروف عدم التأكد؟ وما هى القرارات الصحيحة منها؟ هل يمكن ارتكاب خطأين فى قرار واحد؟

٣ ـ ما هو الخطأ من النوع الأول؟ وما هو الخطأ من النوع الثانى؟ ولماذا ارتبطت المخاطرة بأحدهما دون الآخر؟

٤ ـ ما هى العلاقة بين : كمية المخاطرة، ومستوى الثقة، ومستوى المعنوية، وقوة الاختبار؟

٥ ـ ما هى المتغيرات الثلاثة التى يعتمد عليها اختبار الفرضية؟

٦ ـ عرف : فرضية العدم، والفرضية البديلة، وعلاقتهما بمنطقة القبول والرفض.

٧ ـ يتخذ القرار بناء على قيمتين، ما هما؟ وما مصدراهما؟

٨ ـ يفترض صاحب مؤسسة أن العاملين بمؤسسته يتمتعون بصفة الأمانة، إلا أن المراجعة اليومية لأموال المؤسسة قد أثبتت أن هناك فرقاً فى الحسابات بطريقة يومية، إذ تكون المبالغ الموردة أقل مما يجب بقليل. وهناك شخص واحد فقط هو المسؤول عن الصندوق، مما جعل صاحب المؤسسة فى حاجة لاتخاذ قرار بشأنه، بأن يعتبره غير أمين ويوجه له تهمة بذلك، أو يعتبر أن هذا الفرق الطفيف يأتى دائماً نتيجة خطأ فى الصرف يرتكبه أمين الصندوق. إزاء ذلك :

أ ـ ما هى فرضية العدم؟

ب ـ ما هو الخطأ من النوع الأول والخطأ من النوع الثانى؟
(حدد أنواع القرارات أولاً).

٩ ـ تعتقد إدارة إحدى شركات الخطوط الجوية أن ٧٠٪ من مسافريها يفضلون التدخين داخل الطائرة، فأرادت أن تختبر ذلك. وضح الحالات التى يمكن أن ترتكب فيها أخطاء من النوع الأول والنوع الثانى.

١٠ ـ يرغب أحد المديرين فى تكليف أحد موظفيه بأعباء إدارية إضافية. وضح الحالات التى يمكن أن يرتكب فيها المدير أحد أنواع الأخطاء.

١١ ـ تعتقد إدارة إحدى المؤسسات أن الوسط الحسابى لأعمار العاملين فيها لا يقل عن ٢٨ سنة، وباختيار عينة عشوائية من ٢٠ من العاملين فى تلك المؤسسة اتضح أن أعمارهم كانت كالآتى :

٢٥ ، ٢٨ ، ٣٠ ، ٢٢ ، ٤٢ ، ١٩ ، ٢٠ ، ٢٠ ، ٢٥ ، ٣٠ ، ٦٢ ، ٢٣ ، ١٨ ، ٢٦ ، ٤٤ ، ٦٥ ، ٣٣ ، ٥١ ، ٢٨ ، ١٩

اختبر صحة تلك الفرضية بمستوى معنوية ٥٪ ، إذا علمت أن توزيع أعمار العاملين بتلك المؤسسة طبيعى .

١٢ ـ استخدم بيانات السؤال السابق إذا كان اعتقاد إدارة المؤسسة هو أن الوسط الحسابى للعمر لا يزيد على ٣٥ سنة .

١٣ ـ استخدم بيانات السؤال الحادى عشر إذا كان الاعتقاد هو أن الوسط الحسابى للعمر يساوى ٣٥ سنة .

١٤ ـ تعتقد إدارة أحد المصانع أن التدريب يرفع كفاءة عمال قسم التعبئة البالغ عددهم ٩٠ عاملاً ، وللتأكد من ذلك تم اختيار ٣٦ عاملاً اختياراً عشوائياً ، وتم تدريبهم لفترة معينة ، وأعيدوا بعدها للعمل . اتضح بعد ذلك أن الوسط الحسابى لعدد الصناديق التى قام بتعبئتها المدربون قد بلغ ٩٦ صندوقاً بانحراف معيارى يساوى ٤ ، بينما بلغ الوسط الحسابى لغير المدربين ٧٥ صندوقاً بانحراف معيارى يساوى ١٠ . اختبر بمستوى معنوية ٥٪ ما إذا كان التدريب مفيداً أم لا .

١٥ ـ ترغب إدارة المشتريات بإحدى المؤسسات فى اختيار طلاء مناسب لمبانى المؤسسة ، فلم تستطع ترسية العطاء لأى من شركتين ، فاختارت أربع علب بنفس الحجم من كل نوع . فاتضح أن الوسط الحسابى للنوع الأول ١٧١ متراً مربعاً بانحراف معيارى ١٠ أمتار مربعة ، بينما كان الوسط الحسابى للنوع الثانى ١٦٥ متراً مربعاً بانحراف معيارى ٩ أمتار مربعة . فهل هناك فرق جوهرى بين النوعين بمستوى معنوية ٥٪؟

١٦ ـ أجريت تجربة على عينة عشوائية من ١٠ أشخاص من المصابين بأمراض القلب ، وتعاطوا دواء يعتقد أنه يزيد ضربات القلب فى الدقيقة الواحدة بخمس ضربات ، فكان عدد ضربات القلب لكل مريض قبل وبعد تعاطى الدواء على النحو الآتى :

رقم المريض	قبل الدواء	بعد الدواء
١	٦٣	٧١
٢	٧٠	٧٢
٣	٦٥	٧٢
٤	٦٠	٦٩
٥	٧١	٧٢
٦	٥٨	٦٨
٧	٥٦	٥٦
٨	٦١	٦٣
٩	٥٧	٥٥
١٠	٦٩	٧٢

اختبر صحة الفرضية بمستوى معنوية ٥٪.

١٧ ـ مصنعان لإنتاج أجهزة الهواتف، اختيرت عينة عشوائية من إنتاج المصنع الأول منها قوامها ٥٠٠ جهاز، فاتضح بعد فحصها أن هناك ٥٠ هاتفاً تالفة، بينما اختيرت عينة من إنتاج المصنع الثانى حجمها ٦٠٠ هاتف كان بينها ٤٢ هاتفاً تالفاً. فهل يعتبر إنتاج المصنع الثانى أفضل من الأول بمستوى معنوية ٥٪؟

١٨ ـ اختبر فرضية السؤال السابق بمستوى معنوية ٢,٥٪، لو أن حجم العينة الأولى ٥ هواتف، وحجم العينة الثانية ٨ هواتف، والتالف من النوع الأول هاتف واحد، ومن النوع الثانى هاتفان.

١٩ ـ تعتقد إحدى الشركات الخاصة بإنتاج السيارات أن السيارة تقطع ١٠ كيلومترات بلتر واحد من البنزين بانحراف معيارى يساوى ٣ كيلومترات، إلا أنه، وباختيار عينة قوامها ٢١ سيارة من ذلك النوع، اتضح أن الوسط الحسابى ٩ كيلومترات بانحراف معيارى يساوى ٤ كيلومترات، فهل يعتبر الانحراف المعيارى أكبر من الانحراف المعيارى المحدد بمستوى معنوية ٢,٥٪؟

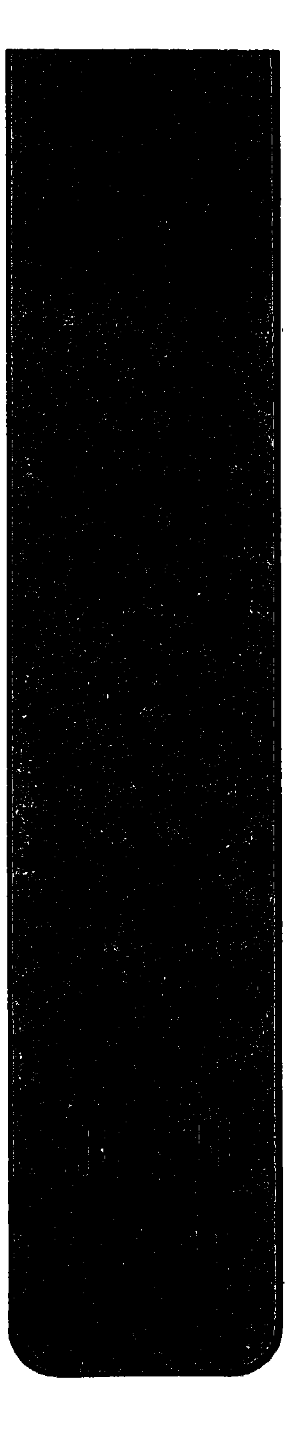
٢٠ ـ كان الانحراف المعيارى لرواتب ٢٠ موظفاً بإحدى الشركات ٢٠٠٠ ريال، بينما كان الانحراف المعيارى لرواتب ١٤ موظفاً بشركة أخرى ١٥٠٠ ريال. فهل يعتبر ذلك دليلاً كافياً بمستوى معنوية ٥٪ على أن الرواتب فى الشركة الثانية أكثر تجانساً؟

٢١ ـ اكتب برنامجاً بلغة بيسك لحساب إحصائية الاختبار للبيانات الواردة بالسؤال (١٤).

٢٢ ـ مستخدماً البيانات الواردة بالسؤال (١٥) ـ اكتب برنامجاً بلغة بيسك لإيجاد إحصائية الاختبار.

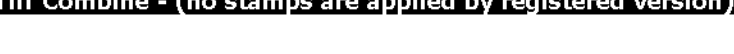
٢٣ ـ استخدم البيانات بالسؤال (١٧)، واكتب برنامجاً بلغة بيسك لإيجاد إحصائية الاختبار.

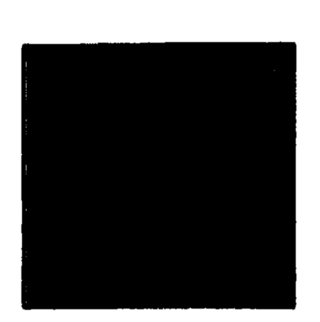




تطبيقات الاختبارات غير المعلمية على البيانات الاسمية والتسلسلية







تطبيقات الاختبارات غير المعلمية على البيانات الاسمية والتسلسلية

١ ـ الفرق بين الاختبارات المعلمية واللامعلمية :

اختصت جميع تطبيقات الفرضيات التى وردت فى الفصل الماضى بمعالم التوزيع الإحصائى ، كالوسط الحسابى والتباين. لذلك فقد اعتمدت جميعها على معرفة التوزيع الإحصائى للمتغيرات ، وتلخص أسلوبها فى مقارنة إحصائية الاختبار (المحسوبة) بالقيمة التى تناظرها فى جدول التوزيع الذى تتبعه تلك الإحصائية (القيمة الحرجة).

إذاً هناك شرطان يجب توفرهما لإجراء تلك الاختبارات ، وهما :

١ ـ معرفة التوزيع الإحصائى للمتغيرات ، وهذا يعنى أن هناك قيداً على التوزيع.

٢ ـ اختبار معالم ذلك التوزيع.

لذلك تسمى تلك التطبيقات بالاختبارات المعلمية (Parametric) . هذا ، ويلاحظ أن مجالات تلك التطبيقات قد انحصرت فى البيانات ذات المقاييس النسبية والمرحلية.

بيد أن هناك حالات لا تكون بياناتها إلا اسمية أو تسلسلية ، كما لا يعرف الشيء الكثير عن التوزيع الإحصائى لتلك البيانات، أو أن هناك علماً مسبقاً بأنه يختلف اختلافاً جوهرياً عن التوزيعات المستخدمة فى التطبيقات المعلمية.

عندئذ لا بد من اتباع أسلوب يخلو من قيد معرفة توزيع المتغير (Distribution – Free Method) ، على ألا يخص الاختبار معلماً من معالم التوزيع المجهول أو تقديره. لذلك فقد أطلق على هذا النوع من الاختبارات الذى يتبع أسلوب التوزيع الحر للمتغيرات ، ولا يهدف لاختبار معالم التوزيع بالاختبارات اللامعلمية (Nonparametric Tests) . هذا ، وتعتبر استمرارية التوزيع الإحصائى هى القيد الوحيد على أسلوب التوزيع الحر للمتغيرات.

وتجدر الإشارة إلى أن هذا التعميم للتوزيعات الإحصائية ، وسهولة تطبيق الاختبارات اللامعلمية ، لا بد أن تكون له بعض السلبيات. وأول تلك السلبيات وأهمها هو أن الاختبارات اللامعلمية لا ترقى لمستوى الاخنبارات المعلمية من حيث الدقة.

بالـرغم من ذلـك فإن مجالات تطبيقـات الاختبارات اللامعلمية كثيرة جداً، خاصة فى المجـالات الاجتـماعية والإدارية، بل لا يمكن حصرها فى فصل واحد؛ لذلك فسوف يتم استعراض أهم تلك الاختبارات وأكثرها استخداماً.

٢ ـ اختبارات البيانات الاسمية (Nominal) :

٢ ـ ١ ماهو المقياس الاسمى؟

المقياس الاسمى للبيانات (كما تم تعريفه من قبل) هو أكثر المقاييس بدائية، إذ يتم تصنيف البيانات حسب صفة أسماء معينة، كالدين، أو الجنس، أو الجنسية أو المهنة .... على ألا يكون هناك تداخل بين المجموعات بعد التصنيف. هذا، ويمكن تقديم أو تأخير أى مجموعة من المجمـوعـات؛ لأن الترتيب لا يعنى التزاماً بالأفضلية. كذلك يمكن استبدال الأسماء بالأرقام (الترميز)، إلا أن القواعد الرياضية لا تطبق على هذه الأرقام؛ إذ لا يمكن جمعها، أو طرحها، أو ضربها، أو قسمتها، وهذا يعنى أن الفروقات بين تلك الرموز لا تعنى شيئاً أيضاً.

٢ ـ ٢ اختبارات حسن المطابقة للبيانات الاسمية لعينة واحدة

: (Goodness – Of Fit Tests)

المقصـود بحسن المـطابقة هو ما إذا كانت القيم العينية (المشاهدات) تطابق (Fit) القيم النظرية أو القيم المتوقعة، فإذا سحبت عينة عشوائية من المصابين بأحد الأمراض الذى يُعتقد أنه لا يخص جنساً دون آخر، فالعدد المتوقع من المصابين من الرجال هو نصف أفراد العينة. أمـا القيمة العينية (المشاهدة) فهو العدد الفعلى للمصابين من بين أفراد العينة. افرض أن حجم العينة كان ٣٠ شخصاً، إذاً فالعدد المتوقع هو ١٥ رجلاً، فإذا اتضح أن هناك ١٩ رجلاً مصاباً بذلك المرض، فهل يعنى ذلك أن الرجال أكثر عرضة لذلك المرض من النساء؟ أم أن هذا الفرق (٤ أشخاص) يمكن أن يعزى للصدفة؛ لأنه غير جوهرى؟ وهنا يأتى دور اختبار حسن المطابقة لعينة واحدة.

افرض أن :

ك ر   تعنى التكرار الفعلى (القيمة العينية) لأى مجموعة.

كَ ر   تعنى التكرار المتوقع (القيمة النظرية) لتلك المجموعة.

د   عدد المجموعات.

ن   حجم العينة.

ففرضية العدم هى :

$ف$: $ك_1 = ك'_1$ ؛ $ك_2 = ك'_2$ ؛ $ك_3 = ك'_3$ ؛ .... ؛ $ك_د = ك'_د$

أما الفرضية البديلة :

$ف_1$: هناك تكرار (ك) واحد على الأقل يختلف اختلافاً جوهرياً عن نظيره المتوقع (ك').

هذا، ويمكن (راجع النظرية التالية إذا دعا الحال) إثبات أن إحصائية الاختبار هى :

$$ك^2 = \frac{\sum_{ر=1}^{د} (ك_ر - ك'_ر)^2}{ك'_ر} \qquad (1)$$

وهى تتبع توزيع مربع كاى على (د - ١) درجات حرية. وعليه تقارن إحصائية الاختبار (المحسوبة) بالقيمة الحرجة (من جدول توزيع مربع كاى بالملحق) بمستوى المعنوية المحدد (أ)، ويكون الفرق جوهرياً إذا كانت إحصائية الاختبار أكبر من القيمة الحرجة.

نظرية (١) :

تعتمد اختبارات حسن المطابقة على ما يسمى بالتوزيع المتعدد الحدود (Multinomial Distribution) وهو امتداد للتوزيع ذى الحدين، الذى ورد ذكره من قبل. فإذا كانت نسبة عناصر النوع الأول هى $ح_1$، ونسبة عناصر النوع الثانى هى $ح_2$، وهكذا حتى أصبحت نسبة عناصر النوع الأخير هى $ح_د$ فإن :

$$\sum_{ر=1}^{د} ح_ر = 1$$

فإذا تم سحب عينة عشوائية حجمها ن عنصراً، فاحتمال أن تتكون من $ن_1$ عنصراً من النوع الأول، و $ن_2$ عنصراً من النوع الثانى، وهكذا حتى أصبح هناك $ن_د$ عنصراً من النوع الأخير هو :

$$\frac{ن!}{ن_1!\ ن_2!\ ن_3! \ldots ن_د!}\ ح_1^{ن_1}\ ح_2^{ن_2}\ ح_3^{ن_3} \ldots ح_د^{ن_د} \qquad (2)$$

نظرية (٢) :

إذا كانت $\text{س}_{١}$ ، $\text{س}_{٢}$ ، $\text{س}_{٣}$ ، . . . . . . ، $\text{س}_{\text{ن}}$ هى متغيرات عشوائية تتبع التوزيع المتعدد، فإن :

$$\text{ك}^{٢} = \frac{\sum_{\text{ر}=١}^{\text{د}} (\text{س}_{\text{ر}} - \text{ن ح}_{\text{ر}})^{٢}}{\text{ن ح}_{\text{ر}}} \qquad (٣)$$

تتبع توزيع مربع كاى على (د ـ ١) درجات حرية، وتزداد دقة هذا التوزيع بزيادة حجم العينة (ن)، كما يعتبر هذا التوزيع مناسباً إذا كان $\text{ن ح}_{\text{ر}} \geqslant ٥$ لكل المجموعات.

افرض أن د = ٢

عندها تصبح

$$\text{س}_{١} + \text{س}_{٢} = \text{ن}$$

$$\text{ح}_{١} + \text{ح}_{٢} = ١$$

$$\therefore \text{س}_{٢} = \text{ن} - \text{س}_{١}$$

$$\text{ح}_{٢} = ١ - \text{ح}_{١}$$

$$\text{ك}^{٢} = \frac{(\text{س}_{١} - \text{ن ح}_{١})^{٢}}{\text{ن ح}_{١}} + \frac{(\text{س}_{٢} - \text{ن ح}_{٢})^{٢}}{\text{ن ح}_{٢}}$$

$$= \frac{(\text{س}_{١} - \text{ن ح}_{١})^{٢}}{\text{ن ح}_{١}} + \frac{(\text{ن} - \text{س}_{١} - \text{ن}(١ - \text{ح}_{١}))^{٢}}{\text{ن}(١ - \text{ح}_{١})}$$

$$= \frac{(\text{س}_{١} - \text{ن ح}_{١})^{٢}}{\text{ن ح}_{١}(١ - \text{ح}_{١})} \qquad (٤)$$

وعليه تكون :

$$\text{ك} = \frac{(\text{س}_{١} - \text{ن ح}_{١})}{\sqrt{\text{ن ح}_{١}(١ - \text{ح}_{١})}}$$

تتبع التوزيع الطبيعى كما ورد فى الفصل الخاص بالتوزيعات. أما مربع التوزيع الطبيعى فهو توزيع مربع كاى على درجة حرية واحدة.

إذاً :

$$ك^{2} = \frac{(س_{1} - ن ح_{1})^{2}}{ن ح_{1} (1 - ح_{1})} \quad (4)$$

تتبع توزيع مربع كاى على درجة حرية واحدة، ومن ثم فإن :

$$ك^{2} = \frac{\sum_{ر=1}^{د} (س_{ر} - ن ح_{ر})^{2}}{ن ح_{1} (1 - ح_{ر})} \quad (5)$$

تتبع توزيع مربع كاى على (د - ١) درجة حرية . وباستبدال :

$$س_{ر} = ك_{ر}$$
$$ن ح_{ر} = ك'_{ر}$$

تصبح

$$ك^{2} = \frac{\sum (ك_{ر} - ك'_{ر})^{2}}{ك'_{ر}} \quad (6)$$

وذلك لأن $\frac{ك'_{ر}}{ن} \longrightarrow$ صفر كلما كان حجم العينة كبيراً.

مثال (٩،١) :

اختيرت عينة عشوائية من ٣٠ متسرباً من متدربى البرامج الإعدادية لخريجى الجامعات بمعهد الإدارة العامة بالرياض، وكان الاعتقاد السائد هو أن عدد المتسربين لا يختلف بين برنامج وآخر (١٠ لكل برنامج)، إلا أن القيم العينية كانت على نحو ما هو مبين فى الجدول أدناه :

جدول (١)
توزيع المتسربين
حسب البرامج لأفراد العينة

البرنامج	عدد المتسربين ($ك_{ر}$)	العدد المتوقع ($ك'_{ر}$)
الأنظمة	٩	١٠
الرقابة المالية	٧	١٠
البنكية المتقدم	١٤	١٠
المجموع	٣٠	٣٠

فهل تختلف نسبة المتسربين اختلافاً جوهرياً عن الثلث (١٠)؟ (مستوى المعنوية ٥٪).

الحل :

ف. : $ك_{١} = ١٠$ ؛ $ك_{٢} = ١٠$ ؛ $ك_{٣} = ١٠$

ف$_{١}$: هناك برنامج واحد على الأقل تختلف نسبته عن الثلث.

أ = ٥٪

د = ٣

د ـ ١ = ٢

القيمة الحرجة من توزيع مربع كاى على درجتى حرية وبمستوى معنوية ٥٪ (جدول توزيع مربع كاى بالملحق) تساوى ٥,٩٩ .

أما إحصائية الاختبار فتستخرج من المعادلة :

$$كا^{٢} = \frac{\sum (ك_{ر} - ك'_{ر})^{٢}}{ك'_{ر}} \qquad (٦)$$

$$= \frac{(٩ - ١٠)^{٢}}{١٠} + \frac{(٧ - ١٠)^{٢}}{١٠} + \frac{(١٤ - ١٠)^{٢}}{١٠}$$

$$= ٢,٦$$

وبما أن ٢,٦ < ٥,٩٩ فلا بد من قبول فرضية العدم، بمعنى أنه ليس هناك دليل على وجود فرق جوهرى بين أعداد المنسحبين من البرامج.

البرنامج التالى يقوم باختبار حسن المطابقة حسب ما هو بالمثال (١,٩)، باستخدام إحصائية الاختبار :

حيث

$S = T/K$

$T = \sum (C_i - K)^2$

K = العدد المتوقع

C = التكرار أو العدد الحقيقى

```
10  REM   اختبارات حسن المطابقه
20  REM  GOODNESS OF FIT
30  DIM C(3)
40  K=10
50  FOR I=1 TO 3
60   READ C(I)            REM  القيم العينيه
70   T=T+(C(I)-K)**2
80  NEXT I
90  S=T/K
100 PRINT ,S;'= احصائيه الاختبار'
110 DATA 9,7,14
120 END

                              المخرجات
                              --------

            2.599999   =  احصائيه الاختبار
```

مثال (٢,٩) :

أجريت دراسة قبل عام لمعرفة أفضل الصحف اليومية حسب آراء القراء، وكانت نتيجة الاستقصاء على النحو المبين بالجدول التالى.

الصحيفة	النسبة المئوية لعدد الذين يفضلونها
أ	٪٦
ب	٪٤٠
جـ	٪١٠
س	٪٢
ص	٪٨
و	٪٢٠
هـ	٪١٤
المجموع	٪١٠٠

جدول (٢)

النسب المئوية حسب آراء القراء

أدت تلك الدراسة لأن تسعى كل صحيفة لتحسين مستواها العام، وبعد مرور عام آخر أعيدت الدراسة لمعرفة ما إذا كان هناك تحول فى آراء بعض القراء، فكان عدد الذين يفضلون كل صحيفة من بين أفراد عينة عشوائية حجمها ٤٠٠ قارىء على النحو التالى :

جدول (٣)
عدد الذين يفضلون فى المرحلة الثانية من الاستقصاء

الصحيفة	توزيع القراء
أ	٢٠
ب	١٦٤
جـ	٤٤
س	١٠
ص	٣٢
و	٨٠
هـ	٥٠
المجموع	٤٠٠

فهل هناك تحول فى آراء القراء؟ (مستوى المعنوية ٥٪).

الحل :

ف : ليس هناك تحول فى الآراء.
ف١ : هناك تحول تجاه صحيفة واحدة على الأقل.

أ = ٥٪
د = ٧
د - ١ = ٦

القيمة الحرجة من جدول توزيع مربع كاى، وبمستوى معنوية ٥٪، وعلى ٦ درجات حرية تساوى ١٢,٥٩ .

أما بالنسبة لإحصائية الاختبار فلا توجد قيم متوقعة ، إذ أن النسب هى التى حلت مكانها ، لذلك تستخدم النسب المئوية الواردة بجدول (٢) لتقدير القيم المتوقعة من مجموع أفراد العينة ، كما هو مبين فى الجدول التالى :

جدول (٤)

إحصائية الاختبار

الصحيفة	النسبة المئوية = ن من جدول (٢)	عدد القراء = ك ر من جدول (٣)	$ك'_ر = \frac{٤٠٠ ن}{١٠٠}$	$ك_ر - ك'_ر$	$\frac{(ك_ر - ك'_ر)^٢}{ك'_ر}$
أ	٦٪	٢٠	٢٤	-٤	٠,٦٧٧
ب	٤٠٪	١٦٤	١٦٠	٤	٠,١٠٠
جـ	١٠٪	٤٤	٤٠	٤	٠,٤٠٠
س	٢٪	١٠	٨	٢	٠,٥٠٠
ص	٨٪	٣٢	٣٢	٠	٠
و	٢٠٪	٨٠	٨٠	٠	٠
هـ	١٤٪	٥٠	٥٦	-٦	٠,٦٤٣
المجموع	١٠٠٪	٤٠٠		صفر	٢,٣١

إذاً إحصائية الاختبار

$$ك^٢ = \frac{\sum (ك_ر - ك'_ر)^٢}{ك'_ر} \qquad (٦)$$

$$= ٢,٣١$$

وبما أن إحصائية الاختبار أقل من القيمة المجدولة (١٢,٥٩) ، فليس هناك دليل كافٍ على أن هناك تحولاً فى آراء القراء ، أى لا بد من قبول فرضية العدم .

فيما يلى برنامج لحساب إحصائية الاختبار لحسن المطابقة عندما لا توجد قيم متوقعة حسب البيانات الواردة بالمثال (٩) باستخدام المعادلة :

$$T = \frac{\sum D^2 (I)}{\sum C \ (I)}$$

حيث :

T = إحصائية الاختبار

D = الفرق بين التكرار الحقيقى والمتوقع

C = التكرار المتوقع اعتماداً على النسبة

```
10  REM    اختبارات حسن المطابقه
20  DIM M$(7),A(7),B(7),C(7),D(7)
40  I.=0                 REM  المجموع
50  T=0
80  FOR I=1 TO 7
90   READ M$(I),A(I),B(I)
100  N=N+B(I)
110 NEXT I
120  PRINT USING 150
122  PRINT USING 140
124  PRINT USING 145
126  PRINT USING 147
130  PRINT USING 150
140 :  (كو - كـر)٢  ك - كـر   كـر    كو      مشاهده النسبه
145 :  ________
147 :     كـر
150 : ____________________________________________________
160 FOR I=1 TO 7
170  C(I)=N*A(I)/100
180  D(I)=B(I)-C(I)
190  E=D(I)**2/C(I)
200  T=T+E
210  PRINT USING 220, E,D(I),C(I),B(I),A(I),M$(I)
220 :     #.###        ###       ###    ###      %##       ##
230 NEXT I
235 PRINT
236 PRINT USING 237, T,N
237 :     #.###       0              ###      %100   المجموع
238 PRINT
240 PRINT ,T;'= اذن احصائيه الاختبار'
260 DATA A,6,20,B,40,164,C,10,44,D,2,10,E,8,32,F,20,80,G,14,50
270 END
```

المخرجات

مشاهده	النسبه	كو	كـر	ك - كـر	(كو - كـر)٢ / كـر
A	% 6	20	24	-4	.667
B	%40	164	160	4	.100
C	%10	44	40	4	.400
D	% 2	10	8	2	.500
E	% 8	32	32	0	0.000
F	%20	80	80	0	0.000
G	%14	50	56	-6	.643
المجموع	%100	400		0	2.310

اذن احصائيه الاختبار = 2.309523

٢ ـ ٣ اختبار المجموعات المترابطة للمشاهدات "Cochran Q test" :

تكون البيانات الاسمية هنا فى شكل مجموعات مترابطة لا يقل عددها عن ثلاث مجموعات لكل فرد من أفراد العينة، بمعنى أن هناك ثلاث تجارب، أو أكثر، على نفس الحقل (الفرد) الذى تكون نتائجه مستقلة عن بقية النتائج. ولتوضيح ذلك خذ المثال التالى :

مثال (٣ ، ٩) :

قام ثلاثة مدربين بتقديم موضوع معين بثلاث طرق مختلفة لاثنى عشر متدرباً تم اختيارهم عشوائياً. وكانت الطريقة الأولى عبارة عن محاضرة فقط، والطريقة الثانية محاضرة مع وسيلة تعليمية (تليفزيون)، بينما كانت الطريقة الثالثة عبارة عن فيلم تليفزيونى فقط.

كان الهدف من إجراء تلك التجارب هو تحديد ما إذا كان هناك فرق جوهرى بين الطرق الثلاث أم لا؛ لذلك فقد سئل كل متدرب عن فاعلية كل طريقة، بأن تكون إجابته بنعم إذا كان يعتقد أنها فعالة، وتكون بلا إذا كانت غير فعالة. والجدول (٥) أدناه يبين الإجابات، بعد أن تم ترميزها بحيث يعنى الرقم (١) الإجابة بنعم، والرقم (صفر) الإجابة بلا. أى الفاعلية وعدم الفاعلية على التوالى.

رقم المتدرب	محاضرة	تليفزيون	محاضرة وتليفزيون	المجموع
١	١	٠	٠	١
٢	٠	٠	١	١
٣	١	١	٠	٢
٤	٠	٠	١	١
٥	١	٠	١	٢
٦	١	١	١	٣
٧	٠	١	١	٢
٨	٠	٠	٠	٠
٩	١	١	١	٣
١٠	٠	٠	١	١
١١	٠	٠	٠	٠
١٢	١	٠	١	٢
المجموع	٦	٤	٨	١٨

جدول (٥)

آراء المتدربين حول فاعلية أو عدم فاعلية طرق التدريب المختلفة

الحل :

$\text{ف}_{٠}$: ليس هناك فرق بين الطرق الثلاث من حيث الفاعلية حسب آراء المتدربين.

$\text{ف}_{١}$: هناك فرق جوهرى بين الطرق الثلاث.

(افرض أن مستوى المعنوية (أ) = ٥٪).

يتلخص الأسلوب المتبع لتحديد القيمة المحسوبة (إحصائية الاختبار) فيما يلى :

إذا كانت :

$\text{س}_{\text{ر}}$ = مجموع أى عمود من الأعمدة (عدد الآراء المؤيدة لوسيلة تدريبية واحدة).

$\text{ص}_{\text{ر}}$ = مجموع أى صف من الصفوف (عدد الحالات الإيجابية لكل متدرب).

د = عدد العينات (عدد التجارب أو المجموعات) وهو عدد الطرق المختلفة فى هذا المثال.

فإحصائية الاختبار هى :

$$\text{كا}^{٢} = \frac{(\text{د} - ١)\left[\text{د} \sum \text{س}_{\text{ر}}^{٢} - \left(\sum \text{س}_{\text{ر}}\right)^{٢}\right]}{\text{د} \sum \text{ص}_{\text{ر}} - \sum \text{ص}_{\text{ر}}^{٢}} \qquad (٧)$$

وهى تتبع توزيع مربع كاى على (د ـ ١) درجات حرية. إذاً فالقيمة الحرجة (المجدولة) تستخرج من جدول توزيع مربع كاى بالملحق على (د ـ ١) درجات حرية.

فالقيمة الحرجة فى هذا المثال من جدول توزيع مربع كاى على درجتى (٣ ـ ١) حرية، وبمستوى معنوية ٥٪ تساوى ٥,٩٩.

أما إحصائية الاختبار فهى :

$$\text{كا}^{٢} = \frac{(٣ - ١)\left[٣(٦^{٢} + ٤^{٢} + ٨^{٢}) - (١٨)^{٢}\right]}{٣ \times ١٨ - (١^{٢} + ١^{٢} + ٢^{٢} + \ldots\ldots + ٠^{٢} + ٢^{٢})}$$

$$= \frac{٢ \times ٢٤}{١٦}$$

$$\text{كا}^{٢} = ٣$$

وبما أن القيمة المحسوبة أقل من الحرجة فليس هناك دليل كافٍ لوجود فرق جوهرى بين الطرق الثلاث. هذا، وتجدر الإشارة هنا إلى أن ذلك يعنى الفرق بين جميع الطرق، ولو أن فرضية العدم قد رفضت، وكان هناك دليل على وجود فرق بين طرق التدريب، فذلك لا يعنى أن الفرق يعزى لطريقة ما. فإذا أراد الباحث تحديد أنجع الطرق فعليه استخدام النسب الواردة فى الفصل السابق لاختبار الفرق بين نسبتين.

البرنامج التالى يقوم باختبار المجموعات المترابطة للمشاهدات حسب ما هو وارد بالمثال (٣) السابق.

لاحظ أننا استخدمنا هنا دالة خاصة، وهى دالة SUM فى الأسطر 40, 50, 60 ، وهى دالة غير مستخدمة فى كل الأجهزة، إلا أنه من الممكن كتابة السطرين 40 – 60 كالآتى إذا لم تكن هذه الدالة متوفرة فى جهازك :

```
35  FOR  I  =  1   To 12
40  P   =  P + A   (I)
45  Q   =  Q + B   (I)
50  R   =  R + C   (I)
60  NEXT  I
```

هذا بالطبع بعد استخدام التعليمة LET لوضع القيمة صفر فى المتغيرات R, Q, P
أما باقى البرنامج فإنه لا يتأثر بأى تغيير.
أما إحصائية الاختبار هنا فهى :

$$V = \frac{(D-1) \times (D\,(P_2 + Q_2 + R_2) - T_2)}{DT - E}$$

حيث :

D = عدد التجارب (العينات أو المجموعات)

P, Q, R = مجاميع الأعمدة الثلاثة

T = المجموع الكلى لعدد الحالات الإيجابية

E = M^2

M = مجموع الصف

```
10  REM    اختبار المجموعات المترابطة للمشاهدات
15  REM  COCHRAN  Q-TEST
20  DIM A(12),B(12),C(12),M(12),E(12),S(12)
40  D=3
50  T=0
60  E=0
80  FOR I=1 TO 12
90   READ A(I),B(I),C(I)
100  M(I)=A(I)+B(I)+C(I)
110  T=T+M(I)
120  E=E+M(I)**2
130 NEXT I
140 P=SUM(A)        REM مجموع العمود الاول
150 Q=SUM(B)        REM مجموع العمود الثاني
160 R=SUM(C)        REM مجموع العمود الثالث
170 PRINT USING 210
180 PRINT USING 200
190 PRINT USING 210
200 :     مجموع    وسيلة ٣   وسيلة ٢    وسيلة ١         رقم
210 : ____________________________________________________________
220 FOR I=1 TO 12
230  S(I)=M(I)**2
240  PRINT USING 250, M(I),C(I),B(I),A(I),I
250 :     ##         ##         ##         ##          ##
260 NEXT I
265 :     ##         ##         ##         ##     المجموع
270 PRINT USING 210
280 PRINT USING 265,T,R,Q,P
290 V=(D-1)*(D*(P*P+Q*Q+R*R)-T*T)/(D*T-E)
300 PRINT ,V;"= احصائية الاختبار"
310 DATA 1,0,0,0,0,1,1,1,0,0,0,1,1,0,1,1,1,1,0,1,1,0,0,0
320 DATA 1,1,1,0,0,1,0,0,0,1,0,1
330 END
```

المخرجات

مجموع	وسيلة ٣	وسيلة ٢	وسيلة ١	رقم
1	0	0	1	1
1	1	0	0	2
2	0	1	1	3
1	1	0	0	4
2	1	0	1	5
3	1	1	1	6
2	1	1	0	7
0	0	0	0	8
3	1	1	1	9
1	1	0	0	10
0	0	0	0	11
2	1	0	1	12
18	8	4	6	المجموع

احصائية الاختبار = 3

٣ ـ اختبارات البيانات التسلسلية (ORDINAL) :

٣ ـ ١ ما هو المقياس التسلسلى؟

المقياس التسلسلى يأتى فى مرحلة أعلى من المقياس الاسمى، إذ يتم تقسيم البيانات (كما ورد من قبل) إلى مجموعات متدرجة وبطريقة مرتبة تصاعدياً أو تنازلياً، إلا أن الفروقات بين كل مجموعتين متتاليتين قد لا تكون متساوية؛ لذلك لا يمكن إخضاعها للعمليات الرياضية. وتعتبر التقديرات الخاصة بدرجات الامتحانات، أو ترتيب نتائج تلاميذ المدارس أمثلة لاستخدامات المقياس التسلسلى (الأول ـ الثانى ـ الثالث. . .). كذلك يعتبر التدرج فى الآراء من الموافقة التامة إلى الرفض مثالاً آخر للبيانات التسلسلية (أوافق جداً، أوافق، لا أدرى، أرفض).

٣ ـ ٢ اختبار حسن المطابقة لعينتين من مجتمعين

(Kolmogorov – Smirnov Test)

المقصود بحسن المطابقة هو تطابق تكرار كل فئة (أو التكرار النسبى الذى يساوى نسبة تكرار الفئة إلى مجموع التكرارات) من العينة الأولى بنظيره فى العينة الثانية.

لقد أثبتت النظريات الإحصائية أن التجمع الصاعد للتكرار النسبى ($\frac{ك_ر}{ن}$) يمثل تقديراً جيداً لتوزيع المجتمع الذى سحبت منه العينة، ويكون الفرق بين التجمعين الصاعدين كبيراً، كلما كان الاختلاف بين التوزيعين كبيراً، والعكس صحيح؛ إذ تعزى الفروقات الطفيفة لعامل الصدفة.

من هذا المنطلق أصبح هذا الاختبار، الخاص بالمقارنة بين عينتين لتحديد ما إذا كان هناك فرق جوهرى، يعتمد على أكبر فرق بين التجمعين الصاعدين للتكرارين النسبيين. هذا، وتنقسم الحالات التطبيقية إلى قسمين اعتماداً على حجم العينة، وسوف يتم استعراض كل حالة على انفراد.

أ ـ عندما يكون حجم كل عينة من العينتين كبيراً :

الحجم الكبير للعينة هنا هو الذى لا يقل عن ٤٠ وحدة، أما طريقة تطبيق هذا الأسلوب فيوضحها المثال التالى :

مثال (٤, ٩) :

أجريت دراسة لقراء إحدى المجلات الكبرى، فاختيرت عينة عشوائية حجمها مائة رجل، وعينة عشوائية أخرى حجمها ثمانون امرأة، وطلب من كل فرد إبداء رأيه فى المستوى العام

للمجلة على مقياس يبتدىء من الصفر إلى ٢٥ ، بحيث يعنى الصفر أضعف مستوى، وتعنى ٢٥ امتيازه، فكانت النتائج على النحو الآتى :

جدول (٦)
آراء أفراد العينة من الجنسين حول المجلة

التقدير (المستوى)	الرجال (ك ر)	النساء (ك َر)
٠ ــ ٥	٣	٥
٥ ــ ١٠	١٤	٣٢
١٠ ــ ١٥	٣١	٢٣
١٥ ــ ٢٠	٤٣	١٤
٢٠ ــ ٢٥	٠٩	٠٦
المجموع	١٠	٨٠

اختبر فرضية العدم :

ف. : ليس هناك فرق جوهرى بين آراء النساء والرجال حول هذه المجلة .

ضد الفرضية البديلة :

ف١ : تقدير الرجال لمستوى هذه المجلة أعلى من تقدير النساء له ـ أى للمستوى، مستوى المعنوية يساوى ٥٪ .

إحصائية الاختبار

إذا كانت :

ل = أكبر فرق بين التكرارين النسبيين التراكميين .

$ن_١$ = حجم العينة الأولى .

$ن_٢$ = حجم العينة الثانية .

فإحصائية الاختبار هى :

$$كا^٢ = \frac{٤ ل^٢ (ن_١ ن_٢)}{ن_١ + ن_٢} \qquad (٨)$$

القيمة الحرجة :

تتبع إحصائية الاختبار السالفة الذكر توزيع مربع كاى على درجتى حرية مهما يكن حجم كل عينة كبيراً. وعليه تكون القيمة الحرجة من جدول توزيع مربع كاى بالملحق على درجتى حرية وبمستوى معنوية ٥٪ تساوى ٥,٩٩ .

بيد أن تحديد قيمة إحصائية الاختبار المبينة فى المعادلة سابقا، يعتمد على أكبر فرق بين التكرارين النسبيين التراكميين (ل). لذلك لا بد من اتباع الخطوات المبينة فى الجدول التالى لاستخراج تلك القيمة.

جدول (٧)

التكرار النسبى والتكرار النسبى التراكمى

التقدير	الرجال ك ر	النساء كَ ر	التكرار النسبى للرجال	التكرار النسبى للنساء	التراكمى رجال	التراكمى نساء	الفرق
٠-٥	٣	٥	٠,٠٣	٠,٠٦	٠,٠٣	٠,٠٦	٠,٠٣
٥-١٠	١٤	٣٢	٠,١٤	٠,٤٠	٠,١٧	٠,٤٦	٠,٢٩٣
١٠-١٥	٣١	٢٣	٠,٣١	٠,٢٩	٠,٤٨	٠,٧٥	٠,٢٧
١٥-٢٠	٤٣	١٤	٠,٤٣	٠,١٧٥	٠,٩١	٠,٩٢٥	٠,٠١٥
٢٠-٢٥	٩	٦	٠,٠٩	٠,٠٧٥	١,٠٠	١,٠٠٠	صفر
المجموع	١٠٠	٨٠	١,٠٠	١,٠٠٠			

وعليه تكون :

ل = ٠,٢٩ (أكبر فرق بين التجمعين التراكميين).

$ن_1$ = ١٠٠ حجم العينة الأولى (رجال).

$ن_2$ = ٨٠ حجم العينة الثانية (نساء).

وبتطبيق معادلة إحصائية الاختبار السالفة الذكر :

$$ك^2 = \frac{٤ \, ل^2 \, ن_1 \, ن_2}{ن_1 + ن_2} \qquad (٨)$$

تصبح

$$ك^{٢} = \frac{٤ \times (٠,٢٩٣)^{٢} \times ١٠٠ \times ٨٠}{١٠٠ + ٨٠}$$

$$= ١٥,٣$$

وبما أن إحصائية الاختبار أكبر من القيمة الحرجة (٥,٩٩) فلا يمكن قبول فرضية العدم. والخلاصة هي أن هناك دليلاً على وجود فرق جوهري بين آراء الرجال والنساء، إذ يمكن القول، وبمستوى معنوية ٥٪، إن الرجال أكثر تفضيلاً لهذه المجلة.

فيما يلي برنامج لحل المثال (٤, ٩) السابق.
نرجو ملاحظة استخدام دالة DSORT فى السطر 190 هذه الدالة مع عبارة MAT تقوم بفرز كل القيم فى المصفوفة D تنازلياً ووضعها فى المصفوفة H . أما إذا لم تكن أوامر المصفوفات متوفرة فى الجهاز الذى لديك فيمكنك الرجوع للمثال بالصفحة ٨٧، حيث هنالك فقرة خاصة بهذا النوع من الفرز.

أما إحصائية الاختبار فقد استخدمت المعادلة :

$$K = \frac{4 L^2 N_1 N_2}{N_1 + N_2}$$

حيث :

L = أكبر فرق بين التكرارين النسبيين

$N1$ = حجم العينة الأولى

$N2$ = حجم العينة الثانية

```
10 REM   برنامج لاختبار حسن المطابقة لعينتين من مجتمعين
20 REM  KOLMOGOROV-SMIRNOV TEST
30 DIM A(5),B(5),M(5),F(5),P(5),Q(5),S(5),T(5),D(5),H(5)
40 READ N    REM  عدد الفئات
50 N1=0     REM  SUM OF M
60 N2=0     REM  SUM OF F
70 FOR I=1 TO N
80   READ A(I),B(I),M(I),F(I)  REM حد ادنى ,حد اعلى ,تكرار١،تكرار٢
90   N1=N1+M(I)
100  N2=N2+F(I)
110 NEXT I
120 FOR I=1 TO N
130  P(I)=M(I)/N1
140  Q(I)=F(I)/N2
150 NEXT I
155 S(1)=P(1)
160 T(1)=Q(1)
162 D(1)=ABS(S(1)-T(1))
165 FOR I=2 TO N
170  S(I)=S(I-1)+P(I)
175  T(I)=T(I-1)+Q(I)
180  D(I)=ABS(S(I)-T(I))
185 NEXT I
190 MAT H=DSORT(D)  REM SORT MATRIX D IN DESCENDING ORDER
195 L=H(1)  REM  أكبر فرق
200 V=4*L**2*N1*N2/(N1+N2)
210 REM  OUTPUT PART
215 PRINT USING 310
220 PRINT USING 290
230 PRINT USING 300
240 PRINT USING 310
250 FOR I=1 TO N
260  PRINT USING 320,D(I),T(I),S(I),Q(I),P(I),F(I),M(I),B(I),A(I)
270  PRINT
280 NEXT I
290 :  الفرق  تراكمي  تراكمي  تكرار نسبي  تكرار نسبي  نساء  رجال  التقدير
300 :         نساء    رجال    نساء        رجال        كر    كر
310 : ________________________________________________________________
320 : #.###   #.###   #.###   #.###       #.###       ###   ###   ## - ##
340 PRINT USING 350,V
350 :              ##.### = احصائية الاختبار
360 DATA 5,0,5,3,5,5,10,14,32,10,15,31,23,15,20,43,14,20,25,9,6
370 END
```

المخرجات

الفرق	تراكمي نساء	تراكمي رجال	تكرار نسبي نساء	تكرار نسبي رجال	نساء كر	رجال كر	التقدير
.033	.063	.030	.063	.030	5	3	5 - 0
.293	.462	.170	.400	.140	32	14	10 - 5
.270	.750	.480	.287	.310	23	31	15 - 10
.015	.925	.910	.175	.430	14	43	20 - 15
.000	1.000	1.000	.075	.090	6	9	25 - 20

15.210 = احصائية الاختبار

لقد كان الاختبار فى المثال السابق ذا اتجاه واحد؛ لأن الفرضية البديلة قد حددت باتجاه واحـد. أما إذا كان الاختبار ذا اتجاهين؛ بسبب عدم تحديد اتجاه معين للفرضية البديلة، فيستخدم الجدول التالى لاستخراج القيمة الحرجة التى تقارن مباشرة بإحصائية الاختبار التى تساوى أكبر فرق بين التجمعين النسبيين التراكميين (ل)، والذى يتم استخراجه بنفس الطريقة المبينة فى جدول (٧).

جدول (٨)*
القيم الحرجة لاختبار عينتين كبيرتين من مجتمعين عندما تكون الفرضية البديلة ذات اتجاهين

مستوى المعنوية (أ)	القيمة التى تكون بعدها ل دليلاً لرفض فرضية العدم حيث : ل = أكبر فرق بين التجمعين النسبيين التراكميين
٪١٠	١,٢٢ $\left(\frac{ن_1 + ن_2}{ن_1 ن_2}\right)^{1/2}$
٪٥	١,٣٦ $\left(\frac{ن_1 + ن_2}{ن_1 ن_2}\right)^{1/2}$
٪٢٥	١,٤٨ $\left(\frac{ن_1 + ن_2}{ن_1 ن_2}\right)^{1/2}$
٪١	١,٦٣ $\left(\frac{ن_1 + ن_2}{ن_1 ن_2}\right)^{1/2}$
٪٠,٥	١,٧٣ $\left(\frac{ن_1 + ن_2}{ن_1 ن_2}\right)^{1/2}$
٪٠,١	١,٩٥ $\left(\frac{ن_1 + ن_2}{ن_1 ن_2}\right)^{1/2}$

فالقيمة الحرجة فى المثال السابق (باعتبار أن مستوى المعنوية ذات الاتجاهين لا يزال ٥٪) تساوى :

$$١,٣٦ \left(\frac{١٠٠ + ٨٠}{١٠٠ \times ٨٠}\right)^{1/2}$$

$$= ٠,٢٠٤$$

وأما إحصائية الاختبار فهى أكبر فرق بين التجمعين (جدول ٧)

$$\therefore ل = ٠,٢٩$$

وبما أن إحصائية الاختبار أكبر من القيمة الحرجة (٠,٢٠٤) فلا بد من رفض فرضية العدم وقبول الفرضية القائلة بأن هناك فرقاً بين آراء الرجال والنساء.

* المصدر :

Isabel S. Patchett; Statistical Methods for Managers and Administrators; VNR, New York; 1982; Page 358.

ب ـ العينة الصغيرة الحجم :

إذا كان حجم أى واحدة من العينتين أقل من ٤٠ ، وكان الحجمان فى ذات الوقت مختلفين ($ن_١ \neq ن_٢$)، فيستخدم أسلوب العينات الكبيرة الوارد فى الفقرة (أ) السالفة الذكر للاختبار.

أما إذا كان حجما العينتين متساويين ($ن_١ = ن_٢$) وصغيرين فى نفس الوقت (أقل من ٤٠) فإحصائية الاختبار هى :

ل = أكبر فرق بين المتجمعين التكراريين الصاعدين، وأما القيمة الحرجة فتستخرج من الجدول رقم (٩) التالى :

جدول (٩) * القيم الحرجة لاختبارات الفرضيات لعينتين صغيرتين متساويتى الحجم

ن حجم العينة	اتجاه واحد للاختبار		اختبار ذو اتجاهين		ن حجم العينة	اتجاه واحد للاختبار		اختبار ذو اتجاهين	
	أ = ٠,٠٥	أ = ٠,٠١	أ = ٠,٠٥	أ = ٠,٠١		أ = ٠,٠٥	أ = ٠,٠١	أ = ٠,٠٥	أ = ٠,٠١
٣	٣	-	-	-	٢١	٨	١٠	٩	١١
٤	٤	-	٤	-	٢٢	٩	١١	٩	١١
٥	٤	٥	٥	٥	٢٣	٩	١١	١٠	١١
٦	٥	٦	٥	٦	٢٤	٩	١١	١٠	١٢
٧	٥	٦	٦	٦	٢٥	٩	١١	١٠	١٢
٨	٥	٦	٦	٧	٢٦	٩	١١	١٠	١٢
٩	٦	٧	٦	٧	٢٧	٩	١٢	١٠	١٢
١٠	٦	٧	٧	٨	٢٨	١٠	١٢	١١	١٣
١١	٦	٨	٧	٨	٢٩	١٠	١٢	١١	١٣
١٢	٦	٨	٧	٨	٣٠	١٠	١٢	١١	١٣
١٣	٧	٨	٧	٩	٣٥	١١	١٣	١٢	
١٤	٧	٨	٨	٩	٤٠	١١	١٤	١٣	
١٥	٧	٩	٨	٩					
١٦	٧	٩	٨	١٠					
١٧	٨	٩	٨	١٠					
١٨	٨	١٠	٩	١٠					
١٩	٨	١٠	٩	١٠					
٢٠	٨	١٠	٩	١١					

* المصدر :

Robert D. Mason; Statistical Techniques In Business and Economics; Third Edition; 1974; Richard D. Irwin, Homewood; Page (637).

مثال (٥ ، ٩) :

افرض أن حجمى العينتين فى المثال (٤) كانا على النحو الآتى :

جدول (١٠)

توزيع تكرارى لآراء الرجال والنساء حول إحدى المجلات

التقدير	الرجال ك ر	النساء كَ رَ	التجمع الصاعد للرجال	التجمع الصاعد للنساء	الفرق
٠ ـ ٥	٠	٣	٠	٣	٣
٥ ـ ١٠	٤	١٢	٤	١٥	١١
١٠ ـ ١٥	١٠	٨	١٤	٢٣	٩
١٥ ـ ٢٠	١٤	٦	٢٨	٢٩	١
٢٠ ـ ٢٥	٢	١	٣٠	٣٠	٠
المجموع	٣٠	٣٠			

∴ إحصائية الاختبار (ل) = ١١

وإذا كان مستوى المعنوية ٥٪ وكان الاختبار ذا اتجاه واحد، كما هو الحال فى المثال (٤)، فالقيمة الحرجة من جدول (٩)، وعند حجم العينة (ن) الذى يساوى ٣٠، تقابل ١٠.

وبما أن إحصائية الاختبار تفوق القيمة الحرجة بقليل فلا يمكن قبول فرضية العدم. هذا، ويلاحظ أن فرضية العدم ستكون مقبولة لو أن مستوى المعنوية (أ) يساوى ١٪؛ لأن القيمة الحرجة بذلك المستوى وبنفس حجم العينة تساوى ١٢.

فيما يلى برنـامـج لحل المثال (٩,٥) السابق، حيث إحصائية الاختبار هنا هى (L) التى تساوى أكبر فرق بين التجمعين التكراريين.

```
10 REM  برنامج لحساب ل عن طريق الفرق بين التجمعين التكراريين الصاعدين
30 DIM A(5),B(5),M(5),F(5),P(5),Q(5),S(5),T(5),D(5),H(5)
40 READ N    REM  عدد الفئات
50 N1=0      REM  SUM OF M
60 N2=0      REM  SUM OF F
70 FOR I=1 TO N
80   READ A(I),B(I),M(I),F(I)
90   N1=N1+M(I)
100  N2=N2+F(I)
110 NEXT I
120 S(1)=M(1)
130 T(1)=F(1)
140 D(1)=ABS(S(1)-T(1))
150 FOR I=2 TO N
160  S(I)=S(I-1)+M(I)
170  T(I)=T(I-1)+F(I)
180  D(I)=ABS(S(I)-T(I))
185 NEXT I
190 MAT H=DSORT(D)  REM SORT MATRIX D IN DESCENDING ORDER
200 L=H(1)  REM  اكبر فرق
210 REM  OUTPUT PART
215 PRINT USING 310
220 PRINT USING 290
230 PRINT USING 300
240 PRINT USING 310
250 FOR I=1 TO N
260  PRINT USING 320,D(I),T(I),S(I),F(I),M(I),B(I),A(I)
270  PRINT
280 NEXT I
290 :        الفرق   تجمع صاعد        تجمع صاعد     نساء    رجال     المقدر
300 :                نساء             رجال          تكر     تكر
310 :       ----------------------------------------------------------------
320 :        ##      ###              ###           ###     ###      ## - ##
340 PRINT USING 350,L
350 :          احصائية الاختبار = ##.###
360 DATA 5,0,5,0,3,5,10,4,12,10,15,10,8,15,20,14,6,20,25,2,1
370 END
```

المخرجات

المقدر	رجال تكر	نساء تكر	تجمع صاعد رجال	تجمع صاعد نساء	الفرق
5 - 0	0	3	0	3	3
10 - 5	4	12	4	15	11
15 - 10	10	8	14	23	9
20 - 15	14	6	28	29	1
25 - 20	2	1	30	30	0

احصائية الاختبار = 11.000

٣ ـ ٢ اختبار مجموع الرتب لعينتين

: (Rank – Sum test)(Mann – Whitney U test)

إذا كان الهدف من الاختبار يخص الفرق بين متوسط رتب العينة الأولى، ومتوسط رتب العينة الثانية ، فاختبار الاستقلال لعينتين من نفس المجتمع أو مجموعتين بنفس التوزيع التالى ذكره هو الأمثل.

البيانات بجدول (١١) التالى عبارة عن درجات ١٢ طالباً من الصف الثالث بمدرسة قيس فى مادة الرياضيات، تقابلها درجات ١٤ طالباً من نفس الصف لنفس الاختبار بمدرسة زهير.
هذا، ولقد تم اختيار العينتين عشوائياً لاختبار فرضية العدم :
ف : مستوى طلاب مدرسة قيس يماثل مستوى طلاب مدرسة زهير فى مادة الرياضيات .

ضد الفرضية البديلة :

$ف_١$: مستوى طلاب مدرسة زهير أفضل من قيس فى مادة الرياضيات .
مستوى المعنوية يساوى ٠,٠٥

يبدأ اختبار مجموع الرتب بترتيب مفردات العينتين معاً ترتيباً تصاعدياً، إذ تستبدل أقل درجة بالعينتين بالرقم (الرتبة) ١، وتستبدل الدرجة التالية لها بالرتبة ٢، والدرجة الثالثة بالرتبة ٣، وهكذا حتى تصبح أعلى درجة بالرتبة ٢٦ . بمعنى أعم فإن الدرجات فى العينتين تستبدل بأرقام متسلسلة أدناها الرقم ١، وأعلاها $ن_١ + ن_٢$، حيث $ن_١$ هى حجم العينة الأولى، و$ن_٢$ هى حجم العينة الثانية. هذا، وتسمى تلك الأرقام المتسلسلة بالرتب. وبذلك يكون المجموع الكلى للرتب بالعينتين $\frac{(ن_١ + ن_٢)}{٢}(ن_١ + ن_٢ + ١)$.

هذا، ويلاحظ من جدول (١١) التالى أنه فى حالة تساوى أكثر من درجة فرتبة كل مشاهدة هى متوسط الرتب المتتالية التى تقابلها؛ فلقد أحرز ثلاثة طلاب بالمدرستين الدرجة ٧٢ (جدول ١١)، بينما أحرز طالب واحد الدرجة ٧١ استحق عليها الرتبة السادسة. إذاً فرتبة كل واحد من الطلاب الثلاثة هى $\frac{٧ + ٨ + ٩}{٣} = ٨$ وبذلك تكون رتبة الطالب الذى حصل على ٧٣ درجة تساوى ١٠ كما هو مبين بالجدول.

إذا كانت :

$ن_١$ = حجم العينة الأولى.

$ن_٢$ = حجم العينة الثانية.

$جـ_١$ = مجموع رتب العينة الأولى.

$جـ_٢$ = مجموع رتب العينة الثانية.

فالإحصائية الخاصة بمجموع الرتب المعروفة بإحصائية مان ويتنى (Mann– Whitney) هى :

$$س_١ = ن_١ ن_٢ + \frac{ن_١(ن_١ + ١)}{٢} - جـ_١ \quad (٩)$$

أو

$$س_٢ = ن_١ ن_٢ + \frac{ن_٢(ن_٢ + ١)}{٢} - جـ_٢ \quad (١٠)$$

هذا، ومن الممكن استخدام الإحصائية $س_١$ أو الإحصائية $س_٢$، إلا أنه يفضل استخدام الإحصائية التى تعتمد على العينة الصغرى لأنها تكون أكثر أماناً عند الاختبار.

الرقم	مدرسة قيس		مدرسة زهير	
	الدرجة	الرتبة	الدرجة	الرتبة
١	٨٣	١٩	٨٦	٢١
٢	٧٢	٨	٧٨	١٥
٣	٧٥	١٢,٥	٧٢	٨
٤	٩٢	٢٥	٦٩	٤
٥	٧٣	١٠	٩٣	٢٦
٦	٧٩	١٦	٧٦	١٤
٧	٦٨	٣	٧٢	٨
٨	٨٧	٢٢	٧٠	٥
٩	٨٤	٢٠	٧٤	١١
١٠	٧٥	١٢,٥	٦٣	٢
١١	٨٢	١٨	٨١	١٧
١٢	٧١	٦	٦١	١
١٣			٨٩	٢٣
١٤			٩١	٢٤
المجموع		١٧٢		١٧٩

جدول (١١)

درجات ١٢ طالباً من مدرسة قيس و١٤ طالباً من مدرسة زهير فى مادة الرياضيات

تتبع إحصائية مان ويتنى السالفة الذكر - التوزيع الطبيعى بوسط حسابى $\bar{س} = \frac{ن_١ ن_٢}{٢}$ وانحراف معيارى

$$ع = \sqrt{\frac{\bar{س}(ن_١ + ن_٢ + ١)}{٦}} \quad (١١)$$

بمعنى أن

$$\text{ع} = \sqrt{\frac{\text{ن}_{١}\,\text{ن}_{٢}\,(\text{ن}_{١} + \text{ن}_{٢} + ١)}{١٢}} \qquad (١٢)$$

وعليه تكون إحصائية الاختبار :

$$\text{ى} = \frac{\text{س} - \bar{\text{س}}}{\text{ع}} \quad \text{تابعة للتوزيع الطبيعى}$$

ويمكن بذلك مقارنة الإحصائية (ى) بالقيمة المجدولة بالملحق (١). أما س فهى $\text{س}_{١}$أو $\text{س}_{٢}$.

هذا، وتجدر الإشارة هنا إلى أن حجم كل عينة من العينتين يجب أن يزيد على ٧ مفردات. أما إذا كان أى من العينتين أقل من ذلك فهناك جداول خاصة بمعالجتها لن نتطرق إليها هنا لندرة استخداماتها.

مثال (٩،٦) :

بمستوى معنوية ٥٪ اختبر فرضية العدم القائلة بأنه لا يوجد فرق جوهرى بين المستوى العام لمدرسة قيس والمستوى العام لمدرسة زهير، ضد الفرضية البديلة القائلة بأن المستوى العام لطلاب مدرسة زهير فى مادة الرياضيات هو الأفضل.

الحل :

$$\text{س}_{١} = \text{ن}_{١}\,\text{ن}_{٢} + \frac{\text{ن}_{١}\,(\text{ن}_{١} + ١)}{٢} - \text{جـ}_{١}$$

$$= ١٢ \times ١٤ + \frac{١٢(١٢ + ١)}{٢} - ١٧٢$$

$$= ٧٤$$

$$\text{س}_{٢} = \text{ن}_{١}\,\text{ن}_{٢} + \frac{\text{ن}_{٢}\,(\text{ن}_{٢} + ١)}{٢} - \text{جـ}_{١}$$

$$= ١٢ \times ١٤ + \frac{١٤\,(١٤ + ١)}{٢} - ١٧٩$$

$$= ٩٤$$

ومن هنا يلاحظ أن :

$$س_١ + س_٢ = ن_١ ن_٢$$

يمكن استخدام الإحصائية $س_١$ أو الإحصائية $س_٢$ إلا أن استخدام الأصغر ($س_١$) يكون أكثر أماناً، وعليه تكون :

$$\bar{س} = \frac{ن_١ ن_٢}{٢}$$

$$= \frac{١٦٨}{٢}$$

$$= ٨٤$$

$$ع = \sqrt{\frac{\bar{س}\,(ن_١ + ن_٢ + ١)}{٦}} \qquad (١١)$$

$$= \sqrt{\frac{٨٤\,(١٢ + ١٤ + ١)}{٦}}$$

$$= ١٩,٤$$

إحصائية الاختبار هى :

$$ى = \frac{س_١ - \bar{س}}{ع} \qquad (١٣)$$

$$= \frac{٧٤ - ٨٤}{١٩,٤}$$

$$= -٠,٥١٥$$

وأمـا القيمـة الحرجة من جدول توزيع (ى) بالملحق (١) بمستوى معنوية ٥٪ فتساوى ١,٩٦ . وبما أن القيمة المطلقة لإحصائية الاختبار (٠,٥١٥) ليست أكبر من القيمة الحرجة فلايمكن رفض فرضية العدم.

البرنامج التالى يقوم باختيار مجموع الرتب لعينتين حسب طريقة مان ويتنى وتستخدم هنا البيانات بالمثال (٦ , ٩).

أما إحصائية الاختيار هنا فهى :

$$W = \frac{D_1 - u}{v}$$

حيث :

$$D_1 = N_1 N_2 + \frac{N_1 (N_1 + 1)}{2} - T_1$$ = س١

N_1 = حجم العينة الأولى

N_2 = حجم العينة الثانية

T_1 = مجموع رتب العينة الأولى (F_1)

$$U = \frac{N_1 N_2}{2}$$ = سَ

$$V = \sqrt{U (N_1 + N_2 + 1)/6}$$ = ع

```
01  REM  برنامج لاختبار مجموع الرتب لعينتين
02  REM  RANK-SUM TEST   BY  MANN-WHITNY U-TEST
10  DIM A(12),B(14),X(26),Y(26),F(12),H(14)
12  T1=0  REM USED TO SUM RANKS OF A
13  T2=0  REM USED TO SUM RANKS OF B
15  READ N1,N2  REM N1=NO OF OBS. FOR A, N2=NO OF OBS. FOR B
20  MAT READ A,B
30  FOR I=1 TO N1
40   X(I)=A(I)
50  NEXT I
60  FOR I=13 TO N1+N2
70   X(I)=B(I-N1)
80  NEXT I
90  MAT A=ASORT(A)  REM SORT MATRIX A
100 MAT B=ASORT(B)  REM SORT MATRIX B
120 MAT X=ASORT(X)  REM SORT MATRIX X
130 C2=1
140 S=1
150 FOR I=2 TO N1+N2
160  IF X(I) = X(I-1) THEN 180 ELSE GOSUB 390
180  S=S+I
190  C2=C2+1
200 NEXT I
210 I=N1+N2+1
220 GOSUB 390
230 FOR I=1 TO N1
240  FOR J=1 TO N1+N2
250   IF A(I)=X(J) THEN 270
260  NEXT J
270  F(I)=Y(J)
280 NEXT I
290 FOR I=1 TO N2
295  FOR J=1 TO N1+N2
300   IF B(I)=X(J) THEN 310
305  NEXT J
310  H(I)=Y(J)
315 NEXT I
316 PRINT USING 360
317 PRINT USING 350
318 PRINT USING 360
319 PRINT
320 FOR I=1 TO N1
325  T1=T1+F(I)
327  T2=T2+H(I)
330  PRINT TAB(7);H(I);TAB(17);B(I);TAB(29);F(I);TAB(38);A(I);TAB(47);I
332 NEXT I
335 FOR I=N1+1 TO N2
340  T2=T2+H(I)
341  PRINT TAB(7);H(I);TAB(17);B(I);TAB(47);I
342 NEXT I
345 PRINT USING 370
346 PRINT USING 371,T2,T1
350 :     الرتبه   مدرسه_ب       الرتبه   مدرسه_ا   مسلسل
360 :  ______________________________________________________
370 :       ___                  ___
371 :      ###.#                ###.#
373 U=(N1*N2)/2 REM ARITHMETIC MEAN
374 V=SQR(U*(N1+N2+1)/6) REM STANDARD DEVIATION
375 D1=N1*N2+(N1*(N1+1))/2-T1
376 W=(D1-U)/V
377 PRINT
378 PRINT USING379,W
379 :             ##.### = احصائيه الاختبار
380 STOP
390 Q=S/C2
400 FOR K=I-C2 TO I-1
410  Y(K)=Q
420 NEXT K
430 C2=0
440 S=0
450 RETURN
460 DATA 12,14,83,72,75,92,73,79,68,87,84,75,82,71
470 DATA 86,78,72,69,93,76,72,70,74,63,81,61,89,91
480 END
```

المخرجات

الرتبه	مدرسه_ب	الرتبه	مدرسه_ا	مسلسل
1	61	3	68	1
2	63	6	71	2
4	69	8	72	3
5	70	10	73	4
8	72	12.5	75	5
8	72	12.5	75	6
11	74	16	79	7
14	76	18	82	8
15	78	19	83	9
17	81	20	84	10
21	86	22	87	11
23	89	25	92	12
24	91			13
26	93			14
179.0		172.0		

احصائيه الاختبار = -0.514

٣ ـ ٤ اختبار مجموع الرتب لأكثر من عينتين (Kruskal – Wallis H – test)

هذا الاختبار عبارة عن امتداد للاختبار السابق، إذ يستخدم لاختبار فرضية العدم القائلة بأنه لا يوجد فرق جوهرى بين أوساط العينات التى يزيد عددها على اثنتين.

لذلك تبدو المرحلة الأولى من مراحل الاختبار بطريقة مشابهة للاختبار السابق، إذ ترتب جميع المتغيرات ترتيباً تصاعدياً، بحيث تستبدل أدنى قيمة بالرتبة ١، وتتدرج تلك الرتب إلى أن تكون رتبة أعلى درجة هى مجموع أحجام العينات.

فإذا كانت :

$ن_1$ = حجم العينة الأولى.

$ن_2$ = حجم العينة الثانية.

$ن_3$ = حجم العينة الثالثة.

⋮

$ن_د$ = حجم العينة الأخيرة

د = عدد العينات وهى ثلاث فأكثر

$$ن = \sum_{ر=1}^{د} ن_ر \qquad (14)$$

فترتيب أعلى قيمة = ن

أما مجموع الرتب فيساوى :

$$\frac{ن}{2}(ن + 1) \qquad (15)$$

وأما إذا كانت :

$جـ_1$ = مجموع رتب العينة الأولى.

$جـ_2$ = مجموع رتب العينة الثانية.

$جـ_3$ = مجموع رتب العينة الثالثة.

⋮

$جـ_د$ = مجموع رتب العينة الأخيرة.

فإحصائية الاختبار هى :

$$كا^2 = \frac{12}{ن(ن + 1)}\left(\sum_{ر=1}^{د} \frac{جـ_ر^2}{ن_ر}\right) - 3(ن + 1) \qquad (16)$$

وهى تتبع توزيع مربع كاى على (د − ١) درجات حرية.

مثال (٩,٧):

اختيرت ثلاث عينات عشوائية من درجات الطلاب الذين تلقوا محاضراتهم فى نفس المادة (الاقتصاد) بثلاث طرق مختلفة ، فكانت درجاتهم على النحو الآتى :

الطريقة / الرقم	محاضرات	تليفزيون	محاضرات وتليفزيون
١	٨٦	٩٣	٨١
٢	٧٩	٧٧	٧٤
٣	٩٥	٨٤	٩٢
٤	٧٣	٨٠	٨٣
٥	٨٢	٩٤	٧١
٦	٨٨	٦٩	٨٩
٧		٩٠	٨٥
٨		٧٢	٩١
٩			٨٧
١٠			٧٨
١١			٧٠

جدول (١٢)
درجات ٣ عينات من الطلاب
فى مادة الاقتصاد

فهل هناك فرق جوهرى بين الطرق الثلاث بمستوى معنوية ٠,٠٥ ؟

الحل :

د = ٣

$ن_{1}$ = ٦

$ن_{2}$ = ٨

$ن_{3}$ = ١١

∴ ن = ٢٥

أ = ٥٪

∴ القيمة الحرجة من جدول توزيع مربع كاى بالملحق (٣) على درجتى حرية (٣ - ١) وبمستوى معنوية ٥٪ تساوى ٥,٩٩١ .

أما بالنسبة لإحصائية الاختبار فتستبدل الدرجات الواردة فى جدول (١٢) بالرتب المبينة فى جدول (١٣) أدناه على النحو الآتى :

الطريقة / الرقم	محاضرات	تليفزيون	محاضرات وتليفزيون
١	١٦	٢٣	١١
٢	٩	٧	٦
٣	٢٥	١٤	٢٢
٤	٥	١٠	١٣
٥	١٢	٢٤	٣
٦	١٨	١	١٩
٧		٢٠	١٥
٨		٤	٢١
٩			١٧
١٠			٨
١١			٢
المجموع	٨٥	١٠٣	١٣٧

جدول (١٣)
بيان برتب الدرجات الواردة
فى جدول (١٢)

$\therefore$ $جـ_{١} = ٨٥$

$جـ_{٢} = ١٠٣$

$جـ_{٣} = ١٣٧$

وعليه تكون إحصائية الاختبار

$$ك^{٢} = \frac{١٢}{ن\,(ن + ١)} \left[\frac{جـ_{١}^{٢}}{ن_{١}} + \frac{جـ_{٢}^{٢}}{ن_{٢}} + \frac{جـ_{٣}^{٢}}{ن_{٣}} \right] - ٣\,(ن + ١)$$

$$= \frac{١٢}{٢٥\,(٢٥ + ١)} \left[\frac{٨٥^{٢}}{٦} + \frac{١٠٣^{٢}}{٨} + \frac{١٣٧^{٢}}{١١} \right] - ٣\,(٢٥ + ١)$$

$$= ٠,٢١٣$$

وبما أن إحصائية الاختبار أقل من القيمة الحرجة (٥,٩٩١)، فلا بد من قبول فرضية العدم طالما أنه ليس هنا دليل كافٍ لرفضها.

فيما يلي برنامج لاختبار مجموع الرتب لأكثر من عينتين بطريقة كروسكال واليس التى سبق شرحها . تستخدم فى هذا البرنامج البيانات المستخدمة فى المثال (٧ , ٩) السابق، وباستخدام إحصائية الاختبار :

$$Z = UV - W$$

حيث :

$$U = \frac{12}{N(N+1)}$$

$$N = N_1 + N_2 + N_3$$

N_1, N_2, N_3 = أحجام العينات

$V = \sum T_{(I)}^{2} / N(I)$ = مجموع مربعات الرتب على الأحجام

$$W = 3(N+1)$$

```
10  REM  برنامج لاختبار مجموع الرتب لاكثر من عينتين
20  REM  KRUSKAL-WALLIS H-TEST
30  DIM A(6),B(8),C(11),X(25),Y(25),F(6),H(8),M(11)
40  T1=0  REM USED TO SUM RANKS OF A
50  T2=0  REM USED TO SUM RANKS OF B
60  T3=0  REM USED TO SUM RANKS OF C
70  READ N1,N2,N3  REM NO OF OBS. FOR A,B,C RESPECTIVELY
80  MAT READ A,B,C
90  FOR I=1 TO N1
100  X(I)=A(I)
110 NEXT I
120 FOR I=N1+1 TO N1+N2
130  X(I)=B(I-N1)
140 NEXT I
150 FOR I=N1+N2+1 TO N1+N2+N3
160  X(I)=C(I-N1-N2)
170 NEXT I
180 MAT A=ASORT(A)  REM SORT MATRIX A
190 MAT B=ASORT(B)  REM SORT MATRIX B
195 MAT C=ASORT(C)  REM SORT MATRIX C
200 MAT X=ASORT(X)  REM SORT MATRIX X
210 C2=1
220 S=1
230 FOR I=2 TO N1+N2+N3
240  IF X(I) = X(I-1) THEN 250 ELSE GOSUB 740
250  S=S+1
260  C2=C2+1
270 NEXT I
280 I=N1+N2+N3+1
290 GOSUB 740
300 FOR I=1 TO N1
310  FOR J=1 TO N1+N2+N3
320   IF A(I)=X(J) THEN 340
330  NEXT J
```

```
340   F(I)=Y(J)
350   T1=T1+F(I)
360  NEXT I
370  FOR I=1 TO N2
380   FOR J=1 TO N1+N2+N3
390    IF B(I)=X(J) THEN 410
400   NEXT J
410   H(I)=Y(J)
420   T2=T2+H(I)
430  NEXT I
440  FOR I=1 TO N3
450   FOR J=1 TO N1+N2+N3
460    IF C(I)=X(J) THEN 480
470   NEXT J
480   M(I)=Y(J)
490   T3=T3+M(I)
500  NEXT I
505  N=N1+N2+N3
510  PRINT USING 630
515  PRINT USING 620
520  PRINT USING 630
530  PRINT
540  FOR I=1 TO N1
541   PRINT TAB(7);M(I);TAB(17);C(I);TAB(26);H(I);TAB(34);B(I);
542   PRINT TAB(44);F(I);TAB(53);A(I);TAB(61);I
543  NEXT I
545  FOR I=N1+1 TO N2
550   PRINT TAB(7);M(I);TAB(17);C(I);TAB(26);H(I);TAB(34);B(I);TAB(61);I
560  NEXT I
565  FOR I=N2+1 TO N3
570   PRINT TAB(7);M(I);TAB(17);C(I);TAB(61);I
580  NEXT I
600  PRINT USING 640
610  PRINT USING 650,T3,T2,T1
620  :    الرتبه   طريقه_٣   الرتبه   طريقه_٢   الرتبه   طريقه_١   مسلسل
630  :     ------------------------------------------------------------
640  :     -----          -----          -----
650  :     ###.#          ###.#          ###.#
660  U=12/(N*(N+1))
661  V=T1**2/N1+T2**2/N2+T3**2/N3
662  W=3*(N+1)
663  Z=U*V-W
670  PRINT
710  PRINT USING 720,Z
720  :             ##.#### = احصائيه الاختبار
730  STOP
740  Q=S/C2
750  FOR K=I-C2 TO I-1
760   Y(K)=Q
770  NEXT K
780  C2=0
790  S=0
800  RETURN
810  DATA 6,8,11,86,79,95,73,82,88,93,77,84,80,94,69,90,72
820  DATA 81,74,92,83,71,89,85,91,87,78,70
830  END
```

المخرجات

مسلسل	طريقه_١	الرتبه	طريقه_٢	الرتبه	طريقه_٣	الرتبه
1	73	5	69	1	70	2
2	79	9	72	4	71	3
3	82	12	77	7	74	6
4	86	16	80	10	78	8
5	88	18	84	14	81	11
6	95	25	90	20	83	13
7			93	23	85	15
8			94	24	87	17
9					89	19
10					91	21
11					92	22
		85.0		103.0		137.0

احصائيه الاختبار = 0.2134

٣ ـ ٥ اختبار فروقات الرتب للأزواج المتقارنة

(Wilcoxon Matched Pair Signed Test)

الاختلاف الأساسى بين العينتين هنا، والعينتين فى اختبار مجموع الرتب الوارد فى (٣-٣)، هو أن العينة الأولى مستقلة عن الثانية فى اختبار مجموع الرتب، بينما تكون العينة الأولى هنا مرتبطة بالثانية بسبب توحيد مصدر البيانات لكل زوج. ذلك لأن التجربتين تجريان على نفس الحقل (الفرد).

فقياس فاعلية الأفلام التليفزيونية مقارنة بالمحاضرات يستدعى أخذ عينتين من المتدربين إذا كان الأسلوب المتبع فى الاختبار هو مجموع الرتب لعينتين، وبعد أن يتم تدريب كل عينة (مجموعة) بمعزل عن المجموعة الأخرى يتم رصد الدرجات الخاصة بالتقييم لمتدربى كل مجموعة. أما إذا اتبعت طريقة اختبار فروقات الرتب للأزواج المتقارنة، فيدرب كل فرد من أفراد عينة واحدة بالطريقتين (الأفلام والمحاضرات)، وترصد الدرجات الخاصة بكل طريقة. والمثال التالى يوضح حالة تطبيقية لاختبار فروقات الرتب.

مثال (٩,٨) :

استجلب قسم للنسخ آلة جديدة؛ لأنها أكثر كفاءة من النوع المستخدم فى ذلك القسم حسب رأى مدير الإدارة التى يتبع لها ذلك القسم، إلا أن رئيس القسم أراد أن يقيس كفاءتها فى السرعة، فاختار عينة عشوائية من ١٦ ناسخاً، ودوَّن سرعة كل منهم على كل من الآلتين فى الدقيقة الواحدة، فحصل على النتائج المبينة بالجدول رقم (١٤) التالى. فهل هناك دليل بمستوى معنوية ٠٥ و٠ على أن الآلة الجديدة أفضل من القديمة ؟

الحل :

يبدو واضحاً أن القرار يعتمد على الفرق بين السرعتين على كل آلة بالنسبة لكل شخص؛ لذلك فقد سُجِّلت تلك الفروقات فى العمود الثالث بالجدول (١٤) التالى.

بيد أن الناسخين ٨، ٩، ١١، ١٣ تساوت سرعاتهم على الآلتين، وأصبح الفرق معدوماً؛ لذلك لابد من استبعادهم من الخطوات التالية؛ لأنهم لا يعطون دليلاً على ميزة أى من الآلتين على الأخرى.

جدول (١٤)

سرعة كل ناسخ من أفراد العينة العشوائية على كل من الآلتين والفرق بين السرعتين.

رقم الناسخ / السرعة في الدقيقة	عدد الكلمات/ الدقيقة على الآلة الحالية	عدد الكلمات/ الدقيقة على الآلة الجديدة	الفرق = الجديدة - الحالية
١	٣١	٤٥	١٤
٢	٦٨	٧٠	٢
٣	٥٣	٥١	٢ -
٤	٤٧	٤٦	١-
٥	٩١	٩٣	٢
٦	٣٦	٣٧	١
٧	٧٢	٨٠	٨
٨	٥١	٥١	صفر
٩	٧٤	٧٤	صفر
١٠	٣٩	٤٩	١٠
١١	٨٦	٨٦	صفر
١٢	٦٢	٧٣	١١
١٣	٤٩	٤٩	صفر
١٤	٨٣	٨٠	٣-
١٥	٧٥	٧٨	٣
١٦	٨١	٨٦	٥

ترتب القيم المطلقة للفروقات (دون اعتبار للإشارة) ترتيباً تصاعدياً بعد استبعاد المشاهدات ذات الفروقات المعدومة، ليصبح عدد المشاهدات ١٢، ثم تحول المراتب إلى رتب على أن تكون إشارة الرتبة هي نفس إشارة الفرق؛ وذلك لمعرفة مجموع الرتب الموجبة، ومجموع الرتب السالبة، كما هو موضح بالجدول (١٥) التالي.

جدول (١٥)

الترتيب والرتب للفروقات

رقم الناسخ	الفروقات	الفروقات المطلقة	ترتيب الفروقات المطلقة	رتب الفروقات المطلقة	الرتب بإشارات الفروقات	
					الموجبة	السالبة
٤	١-	١	١	١,٥		١,٥ -
٦	١	١	٢	١,٥	١,٥	
٢	٢	٢	٣	٤,٠	٤	
٣	٢ -	٢	٤	٤,٠		٤-
٥	٢	٢	٥	٤,٠	٤	
١٤	٣ -	٣	٦	٦,٥		٦,٥ -
١٥	٣	٣	٧	٦,٥	٦,٥	
١٦	٥	٥	٨	٨,٠	٨	
٧	٨	٨	٩	٩,٠	٩	
١٠	١٠	١٠	١٠	١٠,٠	١٠	
١٢	١١	١١	١١	١١,٠	١١	
١	١٤	١٤	١٢	١٢,٠	١٢	
المجموع			٧٨	٧٨	٦٦	١٢ -

يلاحظ من الجدول (١٥) السابق أن رتب الفروقات المتساوية هى متوسط ترتيباتها، كما أن مجموع رتب الفروقات المطلقة هو مجموع ترتيب تلك الفروقات ، فإذا كانت :

$\text{ل}_{١}$ = مجموع الرتب الموجبة .

$\text{ل}_{٢}$ = مجموع الرتب السالبة .

ن = عدد المفردات التى لا تساوى فروقاتها أصفاراً .

فإن :

$$\text{ل}_{١} + |\text{ل}_{٢}| = \frac{\text{ن}}{٢}(\text{ن} + ٢) \qquad (١٧)$$

$$٦٦ + |-١٢| = \frac{١٢}{٢}(١٢ + ١)$$

$$= ٧٨$$

ولو أن الفرق بين الآلتين ضعيف لما اختلف الفرق بين مجموع الرتب الموجبة والسالبة عن الصفر، بمعنى أنهما يقتسمان مجموع الرتب ، أو لا يكون بعد كل منهما عنه كبيراً، إذ تصبح

$$ل_١ = | ل_٢ | = \frac{٧٨}{٢}$$
$$= ٣٩$$

إذاً فإحصائية الاختبار (ل) هى $ل_١$ أو $ل_٢$ أيهما أصغر دون اعتبار للإشارة ويكون حدها الأعلى $\frac{ن}{٤}$(ن + ١)، بمعنى أن :

$$ل \geqslant \frac{ن}{٤} (ن + ١)$$

أما الحد الأدنى لها فهو الذى يبينه جدول (١٦) التالى، وعليه تقبل فرضية العدم إذا كانت القيمة المطلقة لأصغر المجموعين من الرتب (ل) فى الفترة :

$$\left(\text{القيمة المستخرجة من جدول (١٦) التالى}\right) \leqslant ل \leqslant \frac{ن}{٤} (ن + ١) \qquad (١٨)$$

$$\text{فقيمة } \frac{ن}{٤} (ن + ١)$$
$$\text{فى المثال} = \frac{٧٨}{٢}$$
$$= ٣٩$$

أما القيمة المستخرجة من جدول (١٦) عند :

$$ن = ١٢$$
$$أ = ٠,٠٥$$

فتساوى ١٧ لأن الاختبار (الفرضية البديلة) ذو اتجاه واحد. بيد أن إحصائية الاختبار (أصغر المجموعين من الرتب) :

$$ل = ١٢$$

وبما أن إحصائية الاختبار لا تقع ضمن الفترة ١٧ ـــ ٣٩ فلا بد من رفض فرضية العدم، وقبول الفرضية البديلة. أى أن هناك دليلاً بمستوى معنوية ٥٪ على أن الآلة الجديدة أسرع من القديمة؛ لأن الفروقات الموجبة هى الأكثر.

الجدير بالذكر أن الجدول (١٦) لا يشمل جميع أحجام العينات (ن) المتقارنة، فإذا لم يجد القارئ قيمة ن مبوبة بالجدول (١٦) يمكنه استخدام الإحصائية التقريبية التى تتبع التوزيع الطبيعى (ى). وهذه الإحصائية هى :

$$\text{ى} = \frac{(\text{س} - \text{ل} - \frac{١}{٢})}{\text{ع}} \qquad (١٩)$$

حيث :

$$\text{س} = \frac{\text{ن}(\text{ن} + ١)}{٤} \qquad (٢٠)$$

$$\text{ع} = \sqrt{\frac{\text{س}(٢\text{ن} + ١)}{٦}} \qquad (٢١)$$

وبعد حساب إحصائية الاختبار (ى) من المعادلة السابقة تستخرج القيمة الحرجة من جدول التوزيع الطبيعى بالملحق (١).

البرنامج التالى يقوم باختبار فروقات الرتب للأزواج المتقارنة بطريقة ويلكوكسون مستخدماً فى هذا المثال البيانات الواردة بالمثال (٨ , ٩) السابق، حيث إحصائية الاختبار هى (ل) التى تساوى مجموع الرتب الموجبة أو السالبة أيهما أصغر.

```
10  REM  WILCOXON MATCHED PAIR SIGNED TEST
20  REM  اختبار فروقات الرتب للازواج المتقارنه
30  DIM A(16),B(16),C(16),Y(16),D(16),E(16)
40  K=0    REM USED TO COUNT NON-ZERO OBS.
50  L1=0   REM USED TO SUM +VE. RANK DIFFERENCES.
60  L2=0   REM USED TO SUM -VE. RANK DIFFERENCES.
70  S=0    REM USED TO SUM ABS. DIFFERENCES.
80  F=0    REM USED TO SUM RANKS.
90  READ N  REM NO OF PAIRED OBSERVATIONS
100 FOR I=1 TO N
110  READ A(I),B(I)
120  IF A(I)-B(I)=0 THEN 150
130  K=K+1
140  C(K)=A(I)-B(I)
150 NEXT I
160 FOR I=1 TO K-1
170  FOR J=I+1 TO K
180   IF ABS(C(I))<ABS(C(J)) THEN 220
190   T=C(I)
200   C(I)=C(J)
210   C(J)=T
220  NEXT J
230  D(I)=ABS(C(I))
240 NEXT I
250 C2=1
260 S1=1
270 FOR I=2 TO K
280  IF D(I) <> D(I-1) THEN GOSUB 810
290  S1=S1+I
300  C2=C2+1
310 NEXT I
320 I=K+1
330 GOSUB 810
340 FOR I=1 TO K
350  IF C(I)>0 THEN L1=L1+E(I) ELSE L2=L2+E(I)
360 NEXT I
370 REM   OUTPUT PART
380 REM    البيانات الاصليه
385 PRINT USING 490
390 PRINT USING 470
400 PRINT USING 480
410 PRINT USING 490
420 PRINT
430 FOR I=1 TO N
440  PRINT USING 500,A(I)-B(I),B(I),A(I),I
450 NEXT I

460 PRINT
470 :   الفرق        السرعه/دقيقه          السرعه/دقيقه        مسلسل
480 :                الاله الجديده         الاله الحاليه
490 :  ____________________________________________________________
500 :   ###          ##                    ##                  ##
510 REM  البيانات المحسوبه
520 PRINT USING 720
530 PRINT USING 730
540 PRINT USING 740
550 PRINT
560 FOR I=1 TO K
570  PRINT USING 750,E(I),I,D(I),C(I),I
580 S=S+E(I)
590 F=F+I
600 NEXT I
610 PRINT
620 PRINT USING 760
630 PRINT
640 PRINT USING 770,S,F
650 PRINT
660 PRINT USING 780,L1
670 PRINT USING 790,L2
680 PRINT
690 IF ABS(L1)<ABS(L2) THEN PRINT USING 800,ABS(L1) ELSE PRINT USING 800,ABS(L2)
700 PRINT
710 STOP
720 :  رتب الفروقات | ترتيب الفروقات | الفروقات |   الفروقات   |  مسلسل
730 :    المطلقه     |    المطلقه     | المطلقه  |              |
740 : ______________|________________|__________|______________|________
750 :     ##.#            ##             ##           ###          ##
760 :   ______          ____
770 :    ###.#           ###                   المجموع
780 :                          ###  =مجموع الرتب الموجبه
790 :                          ###  =مجموع الرتب السالبه
800 :                          ###  = اذن احصائيه الاختبار
805 REM  RANK SUBROUTINE
810 Q=S1/C2
820 FOR G=I-C2 TO I-1
830  E(G)=Q
840 NEXT G
850 C2=0
860 S1=0
870 RETURN
880 DATA 16,31,45,68,70,53,51,47,46,91,93,36,37,72,80,51,51
890 DATA 74,74,39,49,86,86,62,73,49,49,83,80,75,78,81,86
900 END
```

المخرجات

مسلسل	السرعه/دقيقه الالة الحاليه	السرعه/دقيقه الالة الجديده	الفرق
1	31	45	-14
2	68	70	-2
3	53	51	2
4	47	46	1
5	91	93	-2
6	36	37	-1
7	72	80	-8
8	51	51	0
9	74	74	0
10	39	49	-10
11	86	86	0
12	62	73	-11
13	49	49	0
14	83	80	3
15	75	78	-3
16	81	86	-5

مسلسل	الفروقات	الفروقات المطلقه	ترتيب الفروقات المطلقة	رتب الفروقات المطلقه
1	-1	1	1	1.5
2	1	1	2	1.5
3	-2	2	3	4.0
4	2	2	4	4.0
5	-2	2	5	4.0
6	-3	3	6	6.5
7	3	3	7	6.5
8	-5	5	8	8.0
9	-8	8	9	9.0
10	-10	10	10	10.0
11	-11	11	11	11.0
12	-14	0	12	12.0
المجموع			78	78.0

مجموع الرتب الموجبه= 12

مجموع الرتب السالبه= 66

ادنى احصائيه الاختبار = 12

جدول (١٦) :*

الحدود الدنيا لإحصائيات اختبارات فروقات الرتب للأزواج المتقارنة (ل)

أ = اختباراً باتجاه واحد.

٢ أ = اختباراً باتجاهين.

ن	٢أ=١٥و. أ=٠٧٥و.	٢أ=١٠و. أ=٠٥و.	٢أ=٠٤و. أ=٠٢و.	٢أ=٠٣و. أ=٠١٥و.	٢أ=٠٢و. أ=٠١و.	٢أ=٠,٠١٥ أ=٠,٠٠٧٥	٢أ=٠١و. أ=٠٠٥و.
٤	٠						
٥	١	٠					
٦	٢	٢	٠	٠			
٧	٤	٣	٢	١	٠	٠	
٨	٧	٥	٣	٣	٢	١	٠
٩	٩	٨	٥	٥	٤	٣	١
١٠	١٢	١٠	٨	٧	٦	٥	٣
١١	١٦	١٣	١٠	٩	٨	٧	٥
١٢	١٩	١٧	١٣	١٢	١١	٩	٧
١٣	٢٤	٢١	١٧	١٦	١٤	١٢	٩
١٤	٢٨	٢٥	٢١	١٩	١٨	١٥	١٢
١٥	٣٣	٣٠	٢٥	٢٣	٢١	١٩	١٥
١٦	٣٩	٣٥	٢٩	٢٨	٢٦	٢٣	١٩
١٧	٤٥	٤١	٣٤	٣٣	٣٠	٢٧	٢٣
١٨	٥١	٤٧	٤٠	٣٨	٣٥	٣٢	٢٧
١٩	٥٨	٥٣	٤٦	٤٣	٤١	٣٧	٣٢
٢٠	٦٥	٦٠	٥٢	٥٠	٤٧	٤٣	٣٧
٢١	٧٣	٦٧	٥٨	٥٦	٥٣	٤٩	٤٢
٢٢	٨١	٧٥	٦٥	٦٣	٥٩	٥٥	٤٨
٢٣	٨٩	٨٣	٧٣	٧٠	٦٦	٦٢	٥٤
٢٤	٩٨	٩١	٨١	٧٨	٧٤	٦٩	٦١
٢٥	١٠٥	١٠٠	٨٩	٨٦	٨٢	٧٦	٦٨
٢٦	١١٨	١١٠	٩٨	٩٤	٩٠	٨٤	٧٥
٢٧	١٢٨	١١٩	١٠٧	١٠٣	٩٩	٩٢	٨٣
٢٨	١٣٨	١٣٠	١١٦	١١٢	١٠٨	١٠١	٩١

* المصدر : Isabel S, Patehet, Statistical Methods for Managers and. Administrators, VNR ; New York, 1982 ; page (359)

ن	٢أ=٠٫١٥ أ=٠٫٠٧٥	٢أ=٠٫١٠ أ=٠٫٠٥	٢أ=٠٫٠٤ أ=٠٫٠٢	٢أ=٠٫٠٣ أ=٠٫٠١٥	٢أ=٠٫٠٢ أ=٠٫٠١	٢أ=٠,٠١٥ أ=٠,٠٠٧٥	٢أ=٠٫٠١ أ=٠٫٠٠٥
٢٩	١٥٠	١٤٠	١٢٦	١٢٢	١١٧	١١٠	١٠٠
٣٠	١٦١	١٥١	١٣٧	١٣٢	١٢٧	١٢٠	١٠٩
٣١	١٧٣	١٦٣	١٤٧	١٤٣	١٣٧	١٣٠	١١٨
٣٢	١٨٦	١٧٥	١٥٩	١٥٤	١٤٨	١٤٠	١٢٨
٣٣	١٩٩	١٨٧	١٧٠	١٦٥	١٥٩	١٥١	١٣٨
٣٤	٢١٢	٢٠٠	١٨٢	١٧٧	١٧١	١٦٢	١٤٨
٣٥	٢٢٦	٢١٣	١٩٥	١٨٩	١٨٢	١٧٣	١٥٩
٤٠	٣٠٢	٢٨٦	٢٦٤	٢٥٧	٢٤٩	٢٣٨	٢٢٠
٥٠	٤٨٧	٤٦٦	٤٣٤	٤٢٥	٤١٣	٣٩٧	٣٧٣
٦٠	٧١٨	٦٩٠	٦٤٨	٦٣٦	٦٢٠	٦٠٠	٥٦٧
٧٠	٩٩٥	٩٦٠	٩٠٧	٨٩١	٨٧٢	٨٤٦	٨٠٥
٨٠	١٣١٨	١٢٧٦	١٢١١	١١٩٢	١١٦٨	١١٣٦	١٠٨٦
٩٠	١٦٨٨	١٦٣٨	١٥٦٠	١٥٣٧	١٥٠٩	١٤٧١	١٤١٠
١٠٠	٢١٠٥	٢٠٤٥	١٩٥٥	١٩٢٨	١٨٩٤	١٨٥٠	١٧٧٩

٣ ـ ٦ اختبار التباين لرتب أكثر من عينتين مترابطتين (Friedman Two Way Analysis of Variance):

تكـون البيانـات التسلسلية هنـا فى شكـل مجموعات مترابطة لا يقل عددها عن ثلاث مجموعات لكل مشاهدة، وهى من هذا المفهوم عبارة عن امتداد للبيانات التى يطبق عليها اختبار فروق الرتب للأزواج المتقارنة الوارد فى (٣ ـ ٥).

بيد أنها تبـدو وكأنها فى حالة تشابه تام مع البيانات التى طبق عليها اختبار المجموعات المترابطة للمشاهدات الوارد فى (٢ ـ٣)، إلا أن الاختلاف الرئيسى بين النوعين هو أن اختبار المجموعات المترابطة للمشاهدات لايمكن تطبيقه إلا على البيانات الاسمية التى تصنف على خاصية معينة، كالإجابة بنعم أو لا، فعال أو غير فعال؛ ذكر أو أنثى ؛ ولذلك يتم ترميزها بصفر أو واحد.

أما البيانات التى يطبق عليها هذ الاختبار فتكون على مقياس تسلسلى لمجموعات مترابطة ؛ لأنها تخص نفس الحقل أو الفرد، ولذلك يبدأ هذا الاختبار بترتيب البيانات لكل فرد (حقل)

برتب حسب الأفضلية. وبـذا يكـون الترتيب داخلياً لكل فرد باعتبار أنه مستقل عن بقية الأفراد. والمثال التالى يشبه إلى حد كبير مثال (٣) الوارد فى (٢ - ٣) لتوضيح الفرق بين مجالى تطبيق الاختبارين.

مثال (٩, ٩) :

قام ثلاثة مدربين بتقديم موضوع معين إلى عينة عشوائية حجمها ١٥ متدرباً بعد أن تم تقسيم ذلك الموضوع إلى ثلاث وحدات تدريبية غير متداخلة. هذا، وقد تم تقديم الوحدة الأولى بمحاضرات فقط، بينما كانت وسيلة التدريب للوحدة الثانية هى الأفلام التليفزيونية، واستخدم المدرب الثالث فى تقديم وحدته بعض المحاضرات مع أفلام تليفزيونية.

كانت هناك تقييمات تتم عند نهاية كل وحدة، وكانت الدرجة القصوى هى ١٠ درجات، والجـدول (١٧) التـالى يوضـح تقديرات المتدربين. فهل تختلف الطرق الثلاث من حيث مستوى الفاعلية؟ (مستوى المعنوية ٥٪) .

جدول (١٧) .
التقييمات الخاصة بأفراد العينة البالغ حجمها ١٥ متدرباً

الوسيلة / رقم المتدرب	محاضرات	تليفزيون	محاضرات + تليفزيون
١	٩	٨	١٠
٢	٨	٥	٤
٣	٤	٧	٥
٤	٦	٨	٩
٥	١	٢	٥
٦	٣	صفر	٢
٧	٤	٦	٨
٨	٨	٤	٦
٩	٦	١٠	٧
١٠	٥	٣	٤
١١	٧	٢	٦
١٢	٢	٦	٥
١٣	٥	٤	٧
١٤	٧	٦	٣
١٥	٣	٥	١

الحل :

ف : لايوجد فرق جوهرى بين أوساط الطرق الثلاث.

ف ١ : هناك فرق جوهرى بين أوساط الطرق فيما بينها.

أ = ٠,٠٥

افرض أن :

ن = عدد الوحدات المكونة لكل عينة، وهى تساوى ١٥ فى هذا المثال.

د = عدد العينات المترابطة، وهى تساوى ثلاث عينات فى هذا المثال.

فإذا تم ترتيب درجات كل متدرب ترتيباً تصاعدياً، بحيث تصبح رتبته فى أقل وحدة تساوى واحداً وتساوى اثنين فى الوحدة التالية، وثلاثة فى أعلى وحدة.

جدول (١٨)
الرتب للتقييمات

الوسيلة / رقم المتدرب	محاضرات	تليفزيون	محاضرات + تليفزيون
١	٢	١	٣
٢	٣	٢	١
٣	١	٣	٢
٤	١	٢	٣
٥	١	٢	٣
٦	٣	١	٢
٧	١	٢	٣
٨	٣	١	٢
٩	١	٣	٢
١٠	٣	١	٢
١١	٣	١	٢
١٢	١	٣	٢
١٣	٢	١	٣
١٤	٣	٢	١
١٥	٢	٣	١
المجموع	٣٠	٢٨	٣٢

فإذا كانت :

جـ ر = مجموع رتب العينة، بمعنى أنه فى هذا المثال :

جـ ١ = ٣٠

جـ ٢ = ٢٨

جـ ٣ = ٣٢

فإحصائية الاختبار هى :

$$ك^{٢} = \frac{١٢}{ن د(د + ١)} \sum جـ_{ر}^{٢} - ٣[ن (د + ١)] \qquad (٢٢)$$

وهى تتبع توزيع مربع كاى على (د − ١) درجات حرية. لذلك تستخرج القيمة الحرجة (المجدولة) من جدول توزيع مربع كاى بالملحق (٣) بمستوى المعنوية المحدد، ودرجات حرية أقل من عدد العينات بواحدة .

إذاً فالقيمة الحرجة الخاصة بالمثال السابق تستخرج من جدول توزيع مربع كاى بالملحق رقم (٣) على درجتى حريـة (٣ ـ ١) وبمستوى معنوية ٥٪ وهى تساوى ٥,٩٩ .

أما إحصائية الاختبار فهى :

$$ك^{٢} = \frac{١٢}{١٥ \times ٣ \times ٤} (٣٠^{٢} + ٢٨^{٢} + ٣٢^{٢}) - ٣ \times ١٥ \times ٤$$

$$= ٠,٥٣$$

وبما أن إحصائية الاختبار أقل من القيمة الحرجة، فليس هناك ما يمنع قبول فرضية العدم بمستوى معنوية ٥٪ ، بمعنى أنه لا يوجد دليل على وجود فرق جوهرى بين الثلاث طرق.

البرنامج التالى يقوم باختبار التباين للرتب لأكثر من عينتين مترابطتين، وهو هنا يقوم باختبار البيانات الواردة بالمثال (٩,٩) السابق بإحصائية الاختبار :

$$K = \frac{12}{ND\,(D+I)} \left[P_1^2 \;+R_{,1}^2 \;+Q_1^2\right] -!3 \left[\cdot N(D+1)\right]$$

حيث :

X_1 = وسط العينة الأولى

X_2 = وسط العينة الثانية

B = CA

C = قيمة ت المجدولة على (ن ١ + ن ٢ - ٢) = ٢,٠

$$A = \left[\frac{(N_1-1)V_1+(N_2-1)V_2}{N_1+N_2-2}\right]\left[\frac{1}{N_1}+\frac{1}{N_2}\right]$$

N_1 = حجم العينة الأولى

N_2 = حجم العينة الثانية

V_1 = تباين العينة الأولى

V_2 = تباين العينة الثانية

```
10   REM   برنامج لحساب حدود الثقه للفرق بين وسطين فى حالة
20   REM   تساوى التباينين للمجتمعين مع عدم معرفتهما
30   READ N1,N2,X1,X2,V1,V2,C
40   DATA 25,40,6000,5300,10,8,2
50   V1=V1*V1
60   V2=V2*V2
70   A=(((N1-1)*V1+(N2-1)*V2)/(N1+N2-2))*(1/N1+1/N2)
80   X=ABS(X1-X2)
90   B=C*SQR(A)
100  L=X-B              REM  الحد الادنى
110  H=X+B              REM  الحد الاعلى
120  PRINT TAB(20);L;'= الحد الادنى'
130  PRINT
140  PRINT TAB(20);H;'= الحد الاعلى'
150  PRINT
999  END
```

المخرجات

الحد الادنى = 695.5049

الحد الاعلى = 704.4949

٤ ـ اختبارات الاستقلال بجداول التوافق (Contingency Tables) :

يجىء هذا الأسلوب الهام كآخر موضوع فى الاختبارات المعلمية واللامعلمية ؛ لأنه يستخدم لجميع أنواع البيانات سواء اسمية أو تسلسلية أو مرحلية أو نسبية ، فإذا تم تقسيم المتغيرات إلى عدد من الصفوف (ص) وعدد من الأعمدة (ع)، وفقاً لأسلوبين مختلفين يعتمد كل منهما على متغير خاص به للتصنيف ، فالناتج هو جدول ذو اتجاهين قد تكون صفوفه اسمية وأعمدته تسلسلية مثلاً .

يسمى ذلك الجدول ذو الاتجاهين بجدول التوافق أو جدول الاقتران (Contingency Table) . هذا، ويشترط أن يكون أقل عدد من الصفوف صفين، وأقل عدد من الأعمدة عمودين، ويعبر عنه بأنه جدول توافق ص × ع . وعليه يكون أصغر جدول توافق هو ٢ × ٢ .

يستخدم اختبار الاستقلال بالجدول التوافقي بمستوى معنوية محدد لاختبار فرضية العدم ، التى تنص دائماً على استقلال الصفوف عن الأعمدة. بمعنى أنه لا توجد صلة جوهرية بين صفة التصنيف الأفقى ، وصفة التصنيف الرأسى ، كأن تقول إنه لا توجد علاقة بين التدخين والاصابة بسرطان الرئة إذا قسم أفراد العينة إلى مدخنين وغير مدخنين أفقياً ، وقسموا إلى مصابين وغير مصابين بسرطان الرئة رأسياً، كما هو مبين فى الجدول (٩) التالى .

جدول (١٩)

جدول توافقى لعينة عشوائية من ١٠٠ شخص من عيادة أمراض الصدرية (جدول ٢×٢)

صفة التصنيف الرأسية / صفة التصنيف الأفقية	عدد المصابين بسرطان الرئة	عدد غير المصابين بسرطان الرئة	مجموع الصفوف
عدد الذين يدخنون	٢٢	٢١	٤٣
عدد الذين لايدخنون	٢٥	٣٢	٥٧
مجموع الأعمدة	٤٧	٥٣	١٠٠

وأما الفرضيةالبديلة فتنص على وجود صلة بين الصفتين، بمعنى أن التدخين يزيد من احتمال الاصابة بسرطان الرئة .

لقد ورد فى الباب (٢-٢) الخاص باختبارات حسن المطابقة للبيانات الاسمية لعينة واحدة أنه إذا كانت:

$ك_ر$ = التكرار الفعلى (القيمة العينية) داخل كل خلية (Cell)

$ك'_ر$ = التكرار المتوقع (القيمة النظرية) داخل تلك الخلية.

د = عدد المجموعات أو الخلايا (Cells)

فالإحصائية :

$$كا^{2} = \frac{\sum_{ر=1}^{د} (ك_ر - ك'_ر)^{2}}{ك'_ر} \qquad (٦)$$

تتبع توزيع مربع كاى على (د ـ ١) درجات حرية.

ولعل الإضافة الجديدة هنا هى أن عدد الخلايا (د) أصبح ناتج ضرب عدد الصفوف (ص) فى عدد الأعمدة (ع). أى أن :

د = ص × ع

وعليه تكون إحصائية الاختبار هنا هى :

$$كا^{2} = \frac{\sum_{ر=1}^{د} (ك_ر - ك'_ر)^{2}}{ك'_ر} \qquad (٦)$$

وهى تتبع توزيع مربع كاى على (ص – ١) × (ع – ١) درجات حرية. فإذا كانت إحصائية الاختبار أقل من القيمة المجدولة قُبلت فرضية العدم القائلة بأنه لا توجد صلة بين صفة التصنيف الأفقى والرأسى بمستوى المعنوية المحدد. أما إذا زادت إحصائية الاختبار المحسوبة من الجدول التوافقى على القيمة المجدولة بتوزيع مربع كاى، فسوف ترفض فرضية العدم، وتقبل الفرضية البديلة بالمستوى المحدد.

بيد أن القيمة المتوقعة ($ك'_ر$) تكون دائماً غير معلومة، إلا أنه، بناء على قاعدة الدفاع عن فرضية العدم ما لم يثبت خلاف ذلك؛ فيكون التكرار المتوقع داخل كل خلية فى حالة صحة فرضية العدم هو ناتج ضرب نسبة مجموع تكرارات الصف الذى تتبعه تلك الخلية إلى المجموع الكلى للتكرارات فى مجموع تكرارات العمود الذى تتبعه تلك الخلية.

بمعنى أن :

$$ك'_ر = \frac{\text{مجموع صف الخلية} \times \text{مجموع عمودها}}{\text{المجموع الكلى للتكرارات}} \qquad (٢٣)$$

مثال (٩,١٠) :

أخـذت عينـة عشـوائية من مرضى قسم الباطنية بأحد المستشفيات، وتم تصنيفهم إلى مدخنين وغير مدخنين من جهة، ومصابين وغير مصابين بارتفاع ضغط الدم من جهة أخرى.

فهل يعتبر التدخين سبباً لارتفاع ضغط الدم بمستوى معنوية ٠,٠١ ؟ والبيانات للعينة البالغ حجمها ٥٠ مريضاً هى :

جدول (٢٠)

جدول توافقى ٢×٢ لعينة من مرضى قسم الباطنية بأحد المستشفيات

صفة التصنيف	يدخنون	لا يدخنون	مجموع الصفوف
ضغط الدم مرتفع	١٢	١١	٢٣
ضغط الدم غير مرتفع	٨	١٩	٢٧
مجموع الأعمدة	٢٠	٣٠	٥٠

الحل :

يستخدم جدول التكرارات الفعلية لاستخراج جدول التكرارات المتوقعة باستخدام القاعدة :

$$\text{كَ}_{ر}\text{(المتوقعة)} = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلى}} \qquad (٢٣)$$

والجدول التالى يوضح ذلك.

جدول (٢١)

حساب التكرارات المتوقعة ($\text{كَ}_{ر}$) من جدول (٢٠) الخاص بالتكرارات الفعلية ($\text{ك}_{ر}$)

صفة التصنيف	يدخن	لا يدخن	المجموع
ضغط الدم مرتفع	$\frac{٢٠ \times ٢٣}{٥٠} = ٩,٢$	$\frac{٢٣}{٥٠} \times ٣٠ = ١٣,٨$	٢٣
ضغط الدم غير مرتفع	$\frac{٢٧}{٥٠} \times ٢٠ = ١٠,٨$	$\frac{٢٧}{٥٠} \times ٣٠ = ١٦,٢$	٢٧
المجموع	٢٠	٣٠	٥٠

ويلاحظ أن المجاميع الحدية للتكرارات المتوقعة تساوى المجاميع الحدية للتكرارات الفعلية.

أما إحصائية الاختبار

$$ك^{۲} = \sum_{ر=۱}^{د} \frac{(ك_{ر} - كَ_{ر})^{۲}}{كَ_{ر}} \qquad (٦)$$

فيمكن حسابها من التكرارات بالجدول (٢٠) والجدول (٢١) :

$$\therefore ك^{۲} = \frac{(۱۲ - ۹,۲)^{۲}}{۹,۲} + \frac{(۱۱ - ۱۳,۸)^{۲}}{۱۳,۸} + \frac{(۸ - ۱۰,۸)^{۲}}{۱۰,۸} + \frac{(۱۹ - ۱٦,۲)^{۲}}{۱٦,۲}$$

$$= (۲,۸)^{۲} \left(\frac{۱}{۹,۲} + \frac{۱}{۱۳,۸} + \frac{۱}{۱۰,۸} + \frac{۱}{۱٦,۲} \right)$$

$$\therefore ك^{۲} = ۲,٦۳٤$$

أمـا القيمة الحرجة من جدول توزيع مربع كاى بالملحق وعلى (ص − ١) × (ع − ١) درجات حرية أى على (٢ − ١) × (٢ − ١) = ١ درجة حرية، وبمستوى معنوية ١٪ فتساوى ٦,٦٣.

وبما أن إحصائية الاختبار أقل من القيمة الحرجة، فليس هناك دليل كافٍ بمستوى معنوية ١٪ على وجود علاقة بين التدخين وارتفاع ضغط الدم. بمعنى أنه لا بد من قبول فرضية العدم.

تجدر الإشارة إلى أن هذا الاختبار يمكن أن يؤدى إلى نتائج خاطئة إذا زادت نسبة عدد الخلايا، التى يقل عدد تكراراتها عن خمسة تكرارات، على ٢٠٪ من العدد الكلى للخلايا. لذلك يمكن دمج الخلايا المتقاربة عند الضرورة لتحقيق ذلك الشرط، إلا أنه من الأفضل زيادة حجم العينة إذا كان ذلك ممكناً.

مثال (٩,١١)

البيانات التالية توضح التقديرات النهائية لعينة عشوائية حجمها مائة خريج من خريجى البرامج الإعدادية بمعهد الإدارة العامة. فهل تختلف التقديرات باختلاف البرامج؟ (مستوى المعنوية ١٪).

جدول (٢٢)

التقديرات النهائية لعينة عشوائية من الخريجين حسب البرامج

البرنامج / التقدير	حاسب آلى	دراسات بنكية	إدارة مستشفيات	دراسات مالية	المجموع
ممتاز	١	٢	٤	١	٨
جيد جداً	٤	٥	٩	٥	٢٣
جيد	١٠	١٣	١٥	٦	٤٤
مقبول	٢	٤	٥	٦	١٧
راسب	٥	٢	١	٠	٠٨
المجموع	٢٢	٢٦	٣٤	١٨	١٠٠

عدد الخلايا = ٥×٤ = ٢٠

عدد الخلايا التى يقل عدد تكراراتها عن ٥ تكرارات = ٩ خلايا، وهى أكثر من ٢٠٪ من العدد الكلى للخلايا.

يلاحظ أن هذه التكرارات محصورة بين الامتياز والجيد جداً والمقبول والراسب، لذلك يمكن ضم كل تقديرين متقاربين ليصبح الجدول ٣×٥ ويكون على نحو ما هو مبين فى الجدول (٢٣) الذى يستوفى شروط تطبيق اختبار مربع كاى الخاص بالاستقلال على النحو الآتى:

جدول (٢٣)

التقديرات النهائية لعينة عشوائية لخريجى بعض البرامج بعد دمج بعض الخلايا المتقاربة

البرنامج / التقدير	حاسب آلى	دراسات بنكية	إدارة مستشفيات	دراسات مالية	المجموع
ممتاز / جيد جداً	٥	٧	١٣	٦	٣١
جيد	١٠	١٣	١٥	٦	٤٤
مقبول / راسب	٧	٦	٦	٦	٢٥
المجموع	٢٢	٢٦	٣٤	١٨	١٠٠

ومن ثم يمكن استخراج التكرارات المتوقعة بالمعادلة

$$ك'_{ر} = \frac{\text{مجموع الصف} \times \text{مجموع العمود}}{\text{المجموع الكلى}} \quad (٢٣)$$

جدول (٢٤)

التكرارات المتوقعة للتقديرات النهائية لعينة الخريجين

البرنامج / التقدير	حاسب آلى	دراسات بنكية	إدارة مستشفيات	دراسات مالية	المجموع
ممتاز / جيد جداً	٦,٨٢	٨,٠٦	١٠,٥٤	٥,٥٨	٣١,٠٠
جيد	٩,٦٨	١١,٤٤	١٤,٩٦	٧,٩٢	٤٤,٠٠
مقبول / راسب	٥,٥٠	٦,٥٠	٨,٥٠	٤,٥٠	٢٥,٠٠
المجموع	٢٢,٠٠	٢٦,٠٠	٣٤,٠٠	١٨,٠٠	١٠٠,٠٠

إحصائية الاختبار هى

$$ك^2 = \sum_{ر=1}^{د} \frac{(ك_ر - ك'_ر)^2}{ك'_ر} \qquad (٦)$$

والجدول التالى يبين فروقات $ك_ر - ك'_ر$ والتى يجب أن تكون مجاميعها الحدية أصفاراً.

جدول (٢٥)

($ك_ر - ك'_ر$) من الجدولين (٢٣) و (٢٤)

البرنامج / التقدير	حاسب آلى	دراسات بنكية	إدارة مستشفيات	دراسات مالية	المجموع
ممتاز / جيد جداً	-١,٨٢	-١,٠٦	٢,٤٦	٠,٤٢	صفر
جيد	٠,٣٢	١,٥٦	٠,٠٤	-١,٩٢	صفر
مقبول / راسب	١,٥٠	-٠,٥٠	-٢,٥٠	١,٥٠	صفر
المجموع	صفر	صفر	صفر	صفر	صفر

وعليه تكون إحصائية الاختبار :

$$ك^2 = \frac{(-١,٨٢)^2}{٦,٨٢} + \frac{(-١,٠٦)^2}{٨,٠٦} + \frac{(٢,٤٦)^2}{١٠,٥٤} + ... + \frac{(١,٥)^2}{٤,٥٠}$$

$$= ٣,٦$$

بذلك يكون عدد درجات الحرية لتوزيع مربع كاى بمستوى معنوية ٠,٠١ يساوى

(ص - ١) × (ع - ١)

= (٣ - ١) × (٤ - ١)

= ٢ × ٣

= ٦

إذاً فالقيمة الحرجة المستخلصة من جدول توزيع مربع كاى على ٦ درجات حرية، وبمستوى معنوية ١٪، تساوى ١٦,٨١ .

وبما أن القيمة الحرجة أكبر من إحصائية الاختبار، فليس هناك ما يبرر رفض فرضية العدم بمستوى معنوية ١٪، وعليه فليس هناك دليل كافٍ على أن مستويات التقديرات مرتبطة ببعض البرامج دون الأخرى.

```
10  REM  (1)  اختبارات الاستقلال بجداول التوافق
20  REM  CONTINGENCY TABLES
30  DIM X(5,4),D(5,4),M(5,4),A(5,4),R(5),C(4)
40  T=0   REM USED TO SUM MATRIX ELEMENTS
50  K=0   REM TEST STATISTIC
60  READ N1,N2   REM ROWS,COLUMNS
70  FOR I=1 TO N1
80   FOR J=1 TO N2
90    READ X(I,J)
100   R(I)=R(I)+X(I,J)  REM ROW TOTALS
110   C(J)=C(J)+X(I,J)  REM COLUMN TOTALS
120   T=T+X(I,J)  REM MATRIX TOTAL
130  NEXT J
140 NEXT I
150 PRINT USING 200
160 PRINT USING 210
170 MAT PRINT X
180 PRINT
190 PRINT
200 :                              البيانات الاصليه
210 :
220 L=0
230 FOR I=1 TO N1
240  FOR J=1 TO N2
250   IF X(I,J)<5 THEN L=L+1
260  NEXT J
270 NEXT I
280 P=N1*N2
290 IF L<=P*20/100 THEN 330
300 PRINT '   عدد الخلايا التى يقل عدد تكراراتها عن ٥ اكثر من ٢٠٪ من'
310 PRINT '   العدد الكلى للخلايا .وعليه فلابد من اعاده تصميم الخلايا'
315 PRINT
320 GOTO 999
330 FOR I=1 TO N1
340  FOR J=1 TO N2
350   M(I,J)=R(I)*C(J)/T
360   D(I,J)=X(I,J)-M(I,J)
370   B(I,J)=D(I,J)**2/M(I,J)
380   K=K+B(I,J)
390  NEXT J
400 NEXT I
410 PRINT USING 460
420 PRINT USING 470
430 MAT PRINT M
440 PRINT
450 PRINT
```

التكرارات المتوقعه

```
460 :
470 :
480 PRINT USING 640
490 PRINT USING 650
500 MAT PRINT D
510 PRINT
520 PRINT
```

جدول الفروقات

```
530 :
540 :
550 PRINT ,K;' = احصائيه الاختبار'
560 F=(N1-1)*(N2-1)
570 PRINT
580 PRINT ,F;' = درجات الحريه'
590 PRINT
600 DATA 5,4,1,2,4,1,4,5,9,5,10,13,15,6,2,4,5,6,5,2,1,0
999 END
```

المخرجات

البيانات الاصليه

1	2	4	1
4	5	9	5
10	13	15	6
2	4	5	6
5	2	1	0

عدد الخلايا التى يقل عدد تكراراتها عن ٥ اكثر من ٢٠٪ من العدد الكلى للخلايا .وعليه فلابد من اعاده تصميم الخلايا

```
10  REM  (2)  اختبارات الاستقلال بجداول التوافق
20  REM  CONTINGENCY TABLES
30  DIM X(3,4),D(3,4),M(3,4),A(3,4),R(3),C(4)
40  T=0    REM USED TO SUM MATRIX ELEMENTS
50  K=0    REM TEST STATISTIC
60  READ N1,N2    REM ROWS,COLUMNS
70  FOR I=1 TO N1
80   FOR J=1 TO N2
90    READ X(I,J)
100   R(I)=R(I)+X(I,J)   REM ROW TOTALS
110   C(J)=C(J)+X(I,J)   REM COLUMN TOTALS
120   T=T+X(I,J)   REM MATRIX TOTAL
130  NEXT J
140 NEXT I
150 PRINT USING 200
160 PRINT USING 210
170 MAT PRINT X
180 PRINT
190 PRINT
200 :                              البيانات الاصليـــه
210 :                              -------------------
220 L=0
230 FOR I=1 TO N1
240  FOR J=1 TO N2
250   IF X(I,J)<5 THEN L=L+1
260  NEXT J
270 NEXT I
280 P=N1*N2
290 IF L<=P*20/100 THEN 330
300 PRINT '     عدد الخلايا التى يقل عدد تكراراتها عن ٥ اكثر من ٢٠٪ من'
310 PRINT '     العدد الكلى للخلايا .وعليه فلابد من اعاده تصميم الخلايا'
315 PRINT
320 GOTO 999
330 FOR I=1 TO N1
340  FOR J=1 TO N2
350   M(I,J)=R(I)*C(J)/T
360   D(I,J)=X(I,J)-M(I,J)
370   B(I,J)=D(I,J)**2/M(I,J)
380   K=K+B(I,J)
390  NEXT J
400 NEXT I
410 PRINT USING 460
420 PRINT USING 470
430 MAT PRINT M
440 PRINT
450 PRINT

460 :                              التكرارات المتوقعـــه
470 :                              ---------------------
480 PRINT USING 530
490 PRINT USING 540
500 MAT PRINT D
510 PRINT
520 PRINT
530 :                              جـدول الفـروقـــات
540 :                              ------------------
550 PRINT ,K;' = احصائيه الاختبار'
560 F=(N1-1)*(N2-1)
570 PRINT
580 PRINT ,F;' = درجات الحريه'
590 PRINT
600 DATA 3,4,5,7,13,6,10,13,15,6,7,6,6,6
999 END
```

المخرجات

البيانات الاصليـــه

7	13	6
13	15	6
6	6	6

التكرارات المتوقعه

5.82	8.059999	10.54	5.58
9.679999	11.44	14.96	7.919999
5.5	6.5	8.5	4.5

جدول الفروقات

-1.82	-1.059999	2.46	.4200001
.3200006	1.56	4.000092E-02	-1.919999
1.5	-.5	-2.5	1.5

احصائيه الاختبار = 3.602571

درجات الحريه = 6

تمارين

١ ـ حدد الفرق بين الاختبارات المعلمية واللامعلمية، ومجالات استخدامات كل منها.

٢ ـ ماهي مزايا وعيوب الاختبارات اللامعلمية؟

٣ ـ ماهو المقصود بحسن المطابقة ؟

٤ ـ قام باحث اجتماعي بإجراء دراسة حول البرامج التلفزيونية المفضلة لدى الرجال والنساء، فكانت النتائج العينية على النحو الآتي :

البرنامج / الجنس	الرياضية	الثقافية	الدينية	الأسرة	الأطفال
النساء	٦	٢٤	٣٨	٢٨	٢٤
الرجال	١٠	٣٠	٤٠	٥	١٥

فهل تختلف الأفضلية باختلاف الجنس بمستوى معنوية ٥٪ ؟ .

٥ ـ أرسلت إدارة التخطيط بمعهد الإدارة العامة بالرياض عدداً من الاستبيانات لعينة من خريجي البرامج الإعدادية السابقين، وكان عدد المجيبين وغير المجيبين وتقديراتهم عند التخرج على النحو الآتي :

التقدير	مقبول	جيد	جيدجداً	ممتاز
الذين أجابوا	١٦	٥١	٤٢	٣١
الذين لم يجيبوا	١٢	١٨	١٤	٦

اختـبر الفـرضية القـائلة بأن إعادة الاستبيان مستقلة عن التقدير عند التخرج وذلك بمستوى معنوية ١٪.

٦ ـ ينتج مصنع للعب الأطفال ثلاثة أنواع من اللعب ، وهى : سيارات صغيرة (سيدان)، وشاحنات، وقطارات.

وتنتج تلك اللعب بكميات متساوية باعتبار أن مستوى التسويق للثلاثة أنواع متساوٍ. بيد أن مسحاً للسوق لعينة مكونة من ٣٥٠ قطعة من المبيعات قد أوضح أن عدد السيارات السيدان قد بلغ ١٦٠، وعدد الشاحنات ١١٠، بينما بلغ عدد القطارات ٨٠ فقط . فهل كان اعتقاد الإدارة صحيحاً بمستوى معنوية ٥٪؟

٧ ـ اختبر الفرضية فى السؤال السابق لو أن اعتقاد الإدارة هو أن الاستهلاك ٥٠٪ للسيدان، و٣٠٪ للشاحنات ، ٢٠٪ للقطارات، وكان الإنتاج بهذه النسب.

٨ ـ فى بحث للمشتريات استجوبت عشر إدارات مشتريات كعينة عشوائية بين تلك الإدارات. وكـان المستجوب فى كل إدارة هو مديرها ولقد طلب من المدير الإجابة بنعم أو لا على الطريقة أو الطرق التى يعتقد أنها ملائمة للأجهزة ، علماً بأنه يجوز أن يوافق على أكثر من طريقة من الطرق الثلاث وهى :

أ ـ الطريقة المركزية للشراء.

ب ـ الطريقة غير المركزية للشراء.

جـ ـ الطريقة المرنة (المزدوجة) للشراء.

هذا ويرمز الواحد للإجابة بنعم بينما يرمز الصفر للإجابة بلا ، والإجابات هى :

رقم العينة	مركزية الشراء	لامركزية الشراء	الطريقة المرنة للشراء
١	١	٠	٠
٢	٠	٠	١
٣	٠	١	١
٤	١	٠	٠
٥	٠	٠	١
٦	٠	١	٠
٧	٠	١	٠
٨	٠	٠	١
٩	١	٠	١
١٠	٠	١	١

فهل هناك فرق جوهرى بمستوى ٥٪ يدل على وجود فرق بين الطرق الثلاث حسب آراء المديرين؟ هل هناك فرق جوهرى بين الطريقة المرنة والطريقة المركزية؟.

٩ ـ استجوبت عينة من القيادات العليا، وعينة من التنفيذيين، حول النظام الحالى لتقويم الأداء، فكانت النتائج كالآتى:

الرأى	القياديون	التنفيذيون
غير مقبول	٢	٧
لا أدرى	٠	٦
مقبول	٨	١٠
جيد جداً	١٢	٣
ممتاز	٦	٢

فهل هناك فرق جوهرى بمستوى معنوية ١٠٪ بين آراء القياديين والتنفذيين؟.

١٠ ـ اختبر فرضية السؤال السابق لو كانت الإجابات على النحو الآتى:

الرأى	القياديون	التنفيذيون
غير مقبول	٢٠	٥٠
لا أدرى	٠	٢٠
مقبول	٨٠	١٠٠
جيد جداً	١٢٠	٣٠
ممتاز	٦٠	٢٠

١١ ـ اختبر فرضية السؤال التاسع لو أن الآراء على النحو الآتى :

الرأى	القياديون	التنفيذيون
غير مقبول	٢	٧
لا أدرى	٠	٦
مقبول	٨	٢٣
جيد جداً	٢٢	١٤
ممتاز	٦	٣

١٢ ـ يقوم ثلاثة أساتذة بتدريس مادة واحدة لثلاث مجموعات مختلفة، وبعد انتهاء البرنامج جلس جميع الطلاب فى المجموعات الثلاث لامتحان موحد، فكانت النتائج على النحو الآتى :

الرقم	المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
١	٦٣	٧٢	٥٤
٢	٧٨	٧١	٦٥
٣	٩١	٦٨	٧٣
٤	٥٦	٨١	٨٠
٥	٧٢	٧٠	٨٣
٦	٦٥	٦٢	٧٩
٧	٧٩	٧٥	٧٢
٨	٨٦	٧٤	٦١
٩	٦٤	٨٠	٨٥
١٠	٨٤	٦٧	٥٩
١١	٩٣	٨٧	
١٢	٩٨	٩٦	
١٣		٨٢	
١٤		٨٩	

اختبر فرضية العدم القائلة بأنه لا يوجد فرق جوهرى بين درجات المجموعات الثلاث بمستوى معنوية ٥٪.

١٣ ـ اختبر فرضية العدم فى السؤال السابق للفرق بين المجموعة الأولى والثانية.

١٤ ـ اختبر فرضية العدم فى السؤال الثانى عشر للفرق بين المجموعة الأولى والثالثة.

١٥ ـ اختبر فرضية العدم فى السؤال الثانى عشر للفرق بين المجموعة الثانية والثالثة.

١٦ ـ تنوى إحدى الجهات استحداث نظام التأمين الاجتماعى لعامليها، وكانت هناك ثلاثة مقترحات لنظام التأمين، بينما يقضى الاقتراح الرابع بعدم استخدام أى نظام للتأمين.

أجريت مقابلات لعينة من العاملين فى تلك الجهة لمعرفة ارائهم حول البدائل الأربعة على أن يرصد كل شخص درجة أقصاها ١٠ درجات لأفضل طريقة.

هل هناك فرق بين الطرق الثلاث حسب آراء العاملين بمستوى معنوية ١٠٪؟
والبيانات هى :

رقم العامل	الطريقة الأولى	الطريقة الثانية	الطريقة الثالثة	الطريقة الرابعة (لا تأمين)
١	٦	٨	٤	٠
٢	١٠	٩	٧	٢
٣	٠	٠	٠	١٠
٤	٥	٥	٥	٥
٥	٨	٧	٣	٩
٦	٩	٨	١٠	٧
٧	٧	٧	٨	٠
٨	١٠	٠	٠	٠
٩	٥	٩	٨	٦
١٠	٣	٨	١٠	٤
١١	٠	١٠	٠	٠
١٢	٨	٦	٦	٨
١٣	٩	٨	٧	٩

١٧ ـ كانت إحدى وكالات السفر تنجز معاملاتها يدوياً، إلا أنها رغبة منها فى تطوير عملها جلبت جهازاً للحاسب الآلى، بيد أن بعض العاملين لم يستوعبوا كيفية استخدام الحاسب لإنجاز المعاملات بطريقة أسرع. والعينة العشوائية التالية توضح ذلك لعدد المعاملات اليومية قبل وبعد استخدام الجهاز، فهل هناك زيادة جوهرية فى عدد المعاملات التى تم إنجازها بمستوى ٥٪؟

الرقم	عدد المعاملات قبل الجهاز	عدد المعاملات بعد الجهاز
١	٢٥	٢٨
٢	٢٧	٣٥
٣	٢١	١٧
٤	٢٣	٢٣
٥	٣١	٣٨
٦	٢١	٢٦
٧	٢٩	٢٩
٨	٣٤	٢٨
٩	٢٣	٣٠
١٠	٢٦	٢٦
١١	٢٢	٢٧
١٢	٢٤	٣٠
١٣	٢٥	٢٩
١٤	٢١	٢٧

١٨ ـ استخدام البيانات الواردة بالسؤال (٤)، واكتب برنامج بيسك لإيجاد إحصائية الاختبار.

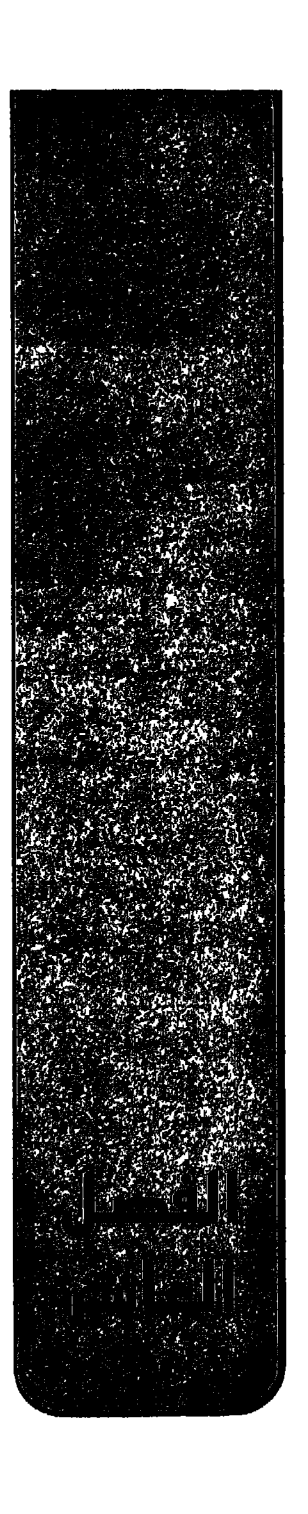
١٩ ـ باستخدام البيانات بالسؤال (٥) اكتب برنامج بيسك لايجاد إحصائية الاختبار.

٢٠ ـ اكتب برنامج بيسك لحساب إحصائية الاختبار للفرضية بالسؤال (٨) مستخدماً نفس البيانات .

٢١ ـ استخدم البيانات بالسؤال (٩) لكتابة برنامج بيسك لحساب إحصائية الاختبار.

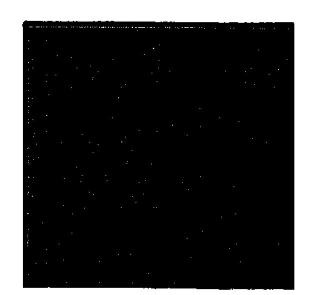
٢٢ ـ اكتب برنامج بيسك لحساب إحصائية الاختبار مستخدماً البيانات الواردة بالسؤال (١٢).

٢٣ ـ اكتب برنامج بيسك لإيجاد إحصائية الاختبار مستخدماً البيانات الواردة بالسؤال (١٦).



الارتباط (CORRELATION)





الارتباط (CORRELATION)

١ - التغاير :

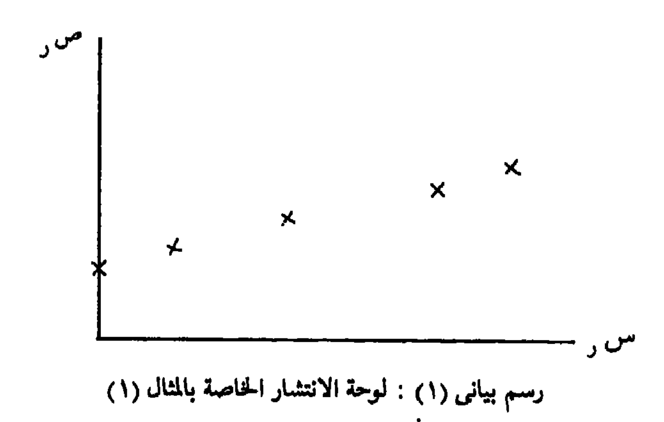
اختصت جميع الحالات السابقة بمعالجة متغير واحد، فالوسط الحسابى، أو الانحراف المعيارى، أو إحصائية الاختبار يخص كلاً منها متغير واحد؛ لأن التوزيع لا يتسع لأكثر من متغير واحد. أما إذا كانت لكل قيمة من المتغير $(س_ر)$ قيمة أخرى تناظرها بمتغير آخر $(ص_ر)$ فالبيانات ذات الأزواج المرتبة تسمى بالبيانات ذات البعدين، ويسمي كل زوج منها بالمتغير العشوائى ذى البعدين. هذا، ويتبع كل متغير من المتغيرين توزيعاً خاصاً يسمى بالتوزيع الهامشى للمتغير.

مثال (١ , ١٠)

البيانات التالية تمثل متغيراً عشوائياً ذا بعدين، هما: الوزن بالأرطال، والعمر بالسنوات، لعينة عشوائية من بعض الصبية المرضى بأحد المستشفيات.

الرقم	الوزن $(س_ر)$	العمر $(ص_ر)$
١	٢٠	٤
٢	٢٦	٧
٣	٣٤	١١
٤	٣٨	١٣
٥	٤٢	١٥
المجموع	١٦٠	٥٠

يلاحظ من المثال (١) أن المتغيرين (س ر ، ص ر) يتزايدان فى اتجاه واحد ، بمعنى أن ص ر توافق س ر فى تغيراتها . والرسم البيانى الذى يسمى لوحة الانتشار (Scatter Diagram) يوضح ذلك .



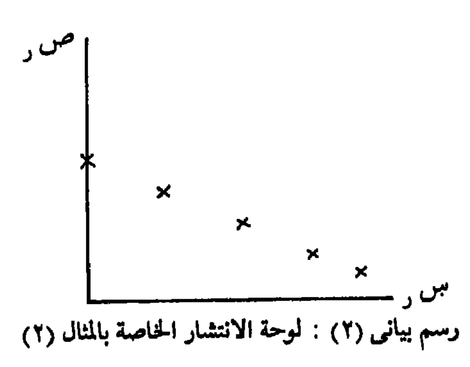
رسم بيانى (١) : لوحة الانتشار الخاصة بالمثال (١)

أما إذا ظلت قيم (س ر) كما كانت عليه بينما بُدِّل اتجاه قيم المتغير (ص ر)، بحيث تقابل أعلى قيمة له لأصغر قيمة للمتغير الأول، وتدرجت بقية القيم على هذا النحو حتى أصبحت أصغر قيمة للمتغير (ص ر) تقابل أكبر قيمة للمتغير (س ر)، مثلما هو موضح فى المثال (٢) التالى، فالمجموع لكل متغير، وكذلك الوسط الحسابى ، والانحراف المعيارى يظل دون أدنى تغيير.

بيد أن اتجاه العلاقة بين المتغيرين قد تبدل تماماً وأصبح فى صورة معاكسة لما هو واضح من لوحة الانتشار السابقة . وهذا ما توضحه لوحة الانتشار بالرسم البيانى رقم (٢) التالى.

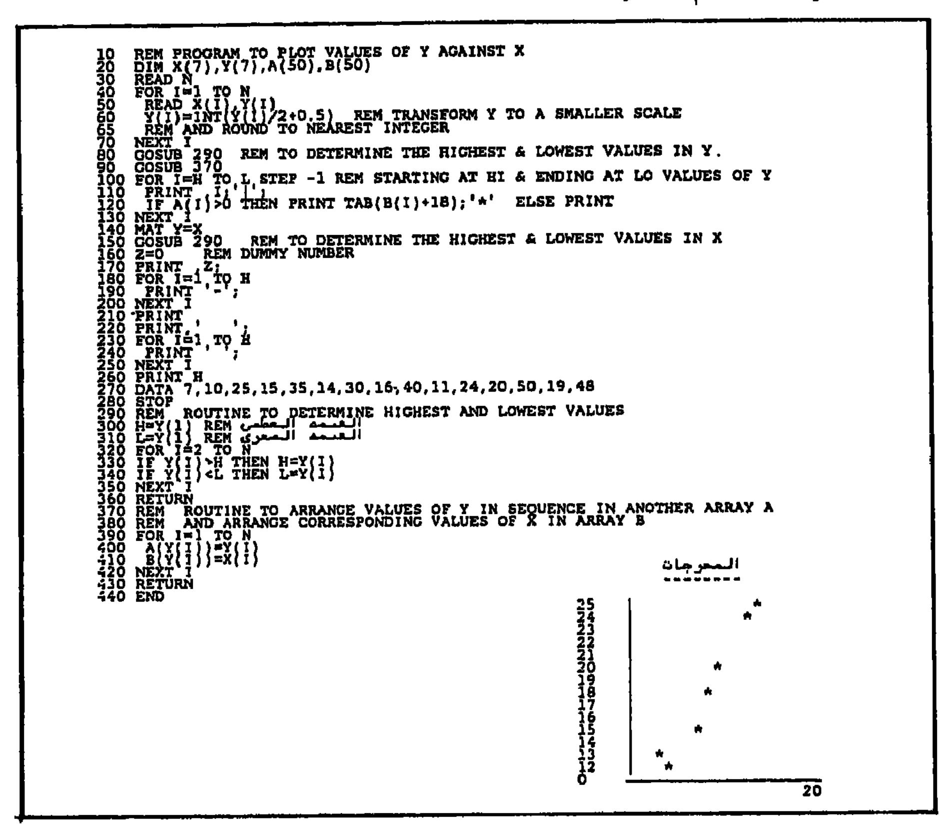
مثال (٢ ، ١٠) :
بيانات باتجاهين متعاكسين

الرقم	س ر	ص ر
١	٢٠	١٥
٢	٢٦	١٣
٣	٣٤	١١
٤	٣٨	٧
٥	٤٢	٤
المجموع	١٦٠	٥٠



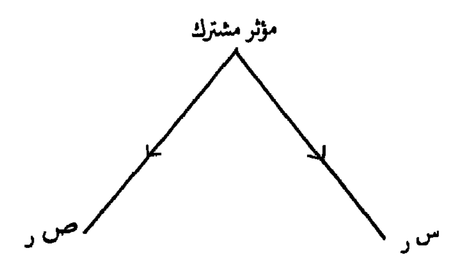
رسم بيانى (٢) : لوحة الانتشار الخاصة بالمثال (٢)

البرنامج التالى مثال لكيفية استخدام عبارات بسيطة بلغة بيسك لرسم بيانى. هنالك بالطبع العديد من الحزم الجاهزة والتى بإمكانها أداء أنواع أكثر صعوبة وتقدماً من المثال الوارد هنا، إلا أن هذا المثال يمكن أن يعطى فكرة عن استخدام دوارة for وعبارة PRINT TAB لتحقيق مثل هذا الرسم البيانى.



```
10  REM PROGRAM TO PLOT VALUES OF Y AGAINST X
20  DIM X(7),Y(7),A(50),B(50)
30  READ N
40  FOR I=1 TO N
50   READ X(I),Y(I)
60   Y(I)=INT(Y(I)/2+0.5)  REM TRANSFORM Y TO A SMALLER SCALE
65   REM AND ROUND TO NEAREST INTEGER
70  NEXT I
80  GOSUB 290  REM TO DETERMINE THE HIGHEST & LOWEST VALUES IN Y.
90  GOSUB 370
100 FOR I=H TO L STEP -1 REM STARTING AT HI & ENDING AT LO VALUES OF Y
110  PRINT ,I;'|';
120  IF A(I)>0 THEN PRINT TAB(B(I)+18);'*'  ELSE PRINT
130 NEXT I
140 MAT Y=X
150 GOSUB 290   REM TO DETERMINE THE HIGHEST & LOWEST VALUES IN X
160 Z=0     REM DUMMY NUMBER
170 PRINT ,Z;
180 FOR I=1 TO H
190  PRINT '-';
200 NEXT I
210 PRINT
220 PRINT,'    ';
230 FOR I=1 TO H
240  PRINT ' ';
250 NEXT I
260 PRINT H
270 DATA 7,10,25,15,35,14,30,16,40,11,24,20,50,19,48
280 STOP
290 REM  ROUTINE TO DETERMINE HIGHEST AND LOWEST VALUES
300 H=Y(1) REM القيمة العظمى
310 L=Y(1) REM القيمة الصغرى
320 FOR I=2 TO N
330 IF Y(I)>H THEN H=Y(I)
340 IF Y(I)<L THEN L=Y(I)
350 NEXT I
360 RETURN
370 REM  ROUTINE TO ARRANGE VALUES OF Y IN SEQUENCE IN ANOTHER ARRAY A
380 REM  AND ARRANGE CORRESPONDING VALUES OF X IN ARRAY B
390 FOR I=1 TO N
400  A(Y(I))=Y(I)
410  B(Y(I))=X(I)
420 NEXT I
430 RETURN
440 END
```

إذاً هناك مؤثر مشترك فى الحالة الأولى، وهو ذو أثر فى اتجاه واحد فى المثال (١)، وفى اتجاهين متعاكسين فى المثال (٢). والشكل التالى يوضح ذلك:



شكل (١) : مؤثر مشترك فى اتجاه واحد

يعتبر التغير حول الوسط الحسابى لكل متغير هو أفضل المقاييس النسبية للزيادة أو النقصان الخاصين بأى متغير. بمعنى أن (س ر - سَ) تمثل تغيرات المتغير الأول، بينما تمثل (ص ر - صَ) التغيرات التى تناظرها فى المتغير الثانى، لذلك تكون (س ر - سَ)(ص ر - صَ) موجبة دائماً إذا كان الأثر فى اتجاه واحد، بينما يكون ناتج الضرب سالبا إذا كان المتغيرات يتأثران بطريقة متعاكسة. وعليه تكون القيمة:

$\sum$ (س ر - سَ) (ص ر - صَ) موجبة إذا كانت العلاقة طردية، وسالبة إذا كانت العلاقة عكسية.

بيد أن القيمة تزيد بزيادة حجم العينة الثنائية (ن). ولتفادى ذلك تقسم القيمة على (ن - ١)، مثلما حدث فى التباين الخاص. بمتغير واحد ليصبح الشكل النهائى لها:

$$\text{ع}_{\text{س ص}} = \frac{\sum (\text{س}_\text{ر} - \bar{\text{س}})(\text{ص}_\text{ر} - \bar{\text{ص}})}{\text{ن} - 1} \quad (1)$$

يلاحظ أن استبدال أى من المتغيرين فى المعادلة رقم (١) يؤدى إلى المعادلة الخاصة بتباين المتغير الآخر، لذلك تسمى الإحصائية $\text{ع}_{\text{س ص}}$ بالتباين المشترك (Covariance) أو التغاير، ويعتبر المصطلح الأخير (التغاير) أكثر شيوعًا فى المراجع العربية. وخلاصة القول أن التغاير العيني هو مجموع مضاريب انحرافات الأزواج العينية المرتبة، مقسوماً على عددها، ناقصاً واحداً.

يكون التغاير موجباً إذا كانت العلاقة طردية، ويكون سالباً إذا كانت العلاقة عكسية. أما إذا كان التغاير معدوماً (صفراً) فهذا دليل على تساوى مجموع مضاريب الانحرافات الموجبة بمجموع مضاريب الانحرافات السالبة. إذاً فليس هناك اتجاه غالب فى هذه الحالة مما يعنى عدم وجود علاقة بين المتغيرين.

وعليه، فإشارة التغاير أهم من قيمته فى هذه المرحلة، إلا إذا كانت صفراً. وبناء على ذلك فالتغاير مقياس نوعى وليس كمياً للعلاقة بين المتغيرين. فالتغاير الكبير لايعنى قوة العلاقة والعكس صحيح، خاصة إذا علمنا أن التغاير يتغير بتغيير وحدة القياس، بيد أنه لا يتأثر بتعـديل نقطة الأصل (الجمع والطرح) وهى خواص مماثلة لخواص التباين. كذلك يمكن عرض المعادلة (١) على النحو التالى :

$$\text{ع}_{\text{س ص}} = \frac{\sum \text{س}_{\text{ر}}\text{ص}_{\text{ر}} - \frac{\sum \text{س}_{\text{ر}} \sum \text{ص}_{\text{ر}}}{\text{ن}}}{\text{ن} - ١} \qquad (٢)$$

هذا، وتجدر الإشارة إلى أن $\text{ع}_{\text{س ص}} = \text{ع}_{\text{ص س}}$ كما هو موضح من المعادلة (٢) السابقة أو المعادلة (١)، كما أنه من السهل إثبات أنه إذا كانت $\text{س}_{\text{ر}} + \text{ص}_{\text{ر}} = \text{ل}_{\text{ر}}$

فإن :

$$\text{ع}^{٢}_{\text{ل}} = \text{ع}^{٢}_{\text{س}} + \text{ع}^{٢}_{\text{ص}} + ٢\,\text{ع}_{\text{س ص}} \qquad (٣)$$

أما إذا كانت :

$$\text{س}_{\text{ر}} - \text{ص}_{\text{ر}} = \text{ل}_{\text{ر}}$$

فإن :

$$\text{ع}^{٢}_{\text{ل}} = \text{ع}^{٢}_{\text{س}} + \text{ع}^{٢}_{\text{ص}} - ٢\,\text{ع}_{\text{س ص}} \qquad (٤)$$

ولهذا التشابه والتداخل بين التغاير والتباين جاء مفهوم مصفوفة التشتت الخاصة بعرض التباينات والتغايرات لعدة متغيرات. ومصفوفة التشتت الخاصة بثلاثة متغيرات س، ص، ل هى :

	س	ص	ل
س	$\text{ع}^{٢}_{\text{س}}$	$\text{ع}_{\text{س ص}}$	$\text{ع}_{\text{س ل}}$
ص	$\text{ع}_{\text{ص س}}$	$\text{ع}^{٢}_{\text{ص}}$	$\text{ع}_{\text{ص ل}}$
ل	$\text{ع}_{\text{ل س}}$	$\text{ع}_{\text{ل ص}}$	$\text{ع}^{٢}_{\text{ل}}$

يلاحظ أن مصفوفة التشتت تكون دائماً متماثلة وعلى قطرها الرئيسى تكـون التباينات وحوله التغايرات، كما أن الاختـلاف بين التغايـر والتباين هو أن التغايـر قد يكـون سالبـاً أو موجبـاً ($-\infty \leq$ ع$_{\text{س ص}}$ $\leq +\infty$).

مثال (٣ , ١٠) :
استخدم البيانات لإيجاد التغاير، ومصفوفة التشتت للمتغيرين س ر ، ص ر ، والبيانات هى :

س ر	ص ر
٢٠	٤
٢٦	٧
٣٤	١١
٣٨	١٣
٤٢	١٥
١٦٠	٥٠

الحل

س ر	ص ر	$\text{س}^2_{\text{ر}}$	$\text{ص}^2_{\text{ر}}$	س ر ص ر
٢٠	٤	٤٠٠	١٦	٨٠
٢٦	٧	٦٧٦	٤٩	١٨٢
٣٤	١١	١١٥٦	١٢١	٣٧٤
٣٨	١٣	١٤٤٤	١٦٩	٤٩٤
٤٢	١٥	١٧٦٤	٢٢٥	٦٣٠
١٦٠	٥٠	٥٤٤٠	٥٨٠	١٧٦٠

$$\text{ع}^2_{\text{س}} = \frac{\sum \text{س}^2_{\text{ر}} - \frac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}}}{\text{ن} - ١}$$

$$= \frac{٥٤٤٠ - \frac{(١٦٠)^2}{٥}}{٤}$$

$$= ٨٠$$

$$ع^{2}_{ص} = \frac{٥٨٠ - \frac{(٥٠)^{2}}{٥}}{٤}$$

$$= ٢٠$$

$$ع_{س ص} = \frac{\sum س_ر ص_ر - \frac{\sum س_ر \sum ص_ر}{ن}}{ن - ١}$$

$$= \frac{١٧٦٠ - \frac{١٦٠ \times ٥٠}{٥}}{٤}$$

$$= \frac{١٦٠}{٤}$$

$$= ٤٠$$

مصفوفة التشتت هى :

$$\begin{bmatrix} ٨٠ & ٤٠ \\ ٤٠ & ٢٠ \end{bmatrix}$$

البرنامج التالى يقوم بحساب التغاير بين متغيرين، والبيانات المستخدمة فى البيانات بالمثال (٣،١٠) السابق باستخدام معادلة التغاير:

$$C = \frac{S - \frac{X_1 Y_1}{N}}{N - 1}$$

حيث :

$$S = \sum X_1 Y_1$$

N = حجم العينة

```
10  REM  PROGRAM TO CALCULATE THE COVARIANCE
30  REM       حساب التغاير بين متغيرين
40  X1=0
50  Y1=0
60  S=0
70  READ N    REM  NO OF OBSERVATIONS
80  FOR I=1 TO N
90   READ X(I),Y(I)   REM المتغيرين
100 NEXT I
110 FOR I=1 TO N
120  X1=X1+X(I)       REM SUM OF X
130  Y1=Y1+Y(I)       REM SUM OF Y
140  S=S+X(I)*Y(I)
150 NEXT I
160 C=(S-(X1*Y1)/N)/(N-1)
170 PRINT 'الوزنXالعمر','العمر','الوزن','الرقم'
180 PRINT '__________','_____','_____','_____'
190 FOR I=1 TO N
200  PRINT X(I)*Y(I),Y(I),X(I),I
210 NEXT I
220 PRINT '_____','______','______'
230 PRINT S,X1,Y1
240 PRINT ,C,' = التغاير'
250 DATA 5,20,4,26,7,34,11,38,13,42,15
260 END
```

المخرجات

الوزنXالعمر	العمر	الوزن	الرقم
80	4	20	1
182	7	26	2
374	11	34	3
494	13	38	4
630	15	42	5
1760	160	50	
	40 = التغاير		

٢ - معامل الارتباط الخطى للبيانات النسبية

(Pearson's Moment Correlation)

يعرف معامل بيرسون العزومى للارتباط الخطى بأنه القيمة المعيارية للتغاير، بمعنى أن معامل الارتباط الخطى هو :

$$\text{ر} = \frac{\text{ع}_{\text{س ص}}}{\text{ع}_{\text{س}}\,\text{ع}_{\text{ص}}} \qquad (٥)$$

$$\text{ر} = \frac{\sum \text{س}_\text{ر}\,\text{ص}_\text{ر} - \dfrac{\sum \text{س}_\text{ر} \sum \text{ص}_\text{ر}}{\text{ن}}}{\sqrt{\left(\sum \text{س}_\text{ر}^{٢} - \dfrac{(\sum \text{س}_\text{ر})^{٢}}{\text{ن}}\right)\left(\sum \text{ص}_\text{ر}^{٢} - \dfrac{(\sum \text{ص}_\text{ر})^{٢}}{\text{ن}}\right)}}$$

هذا، ويلاحظ من المعادلة (٦) أنه إذا كانت :

س ر = ص ر

فإن

ر = ١

أما إذا كانت

س ر = - ص ر

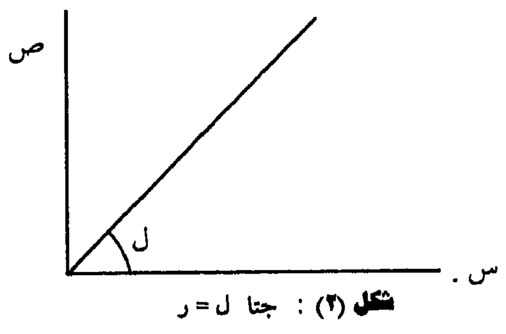
فإن :

ر = - ١

لذلك فللارتباط حدود ؛ إذ أنه يتراوح بين + ١ و - ١ . أى أن :

- ١ ≤ ر ≤ + ١

ويكون الارتباط + ١ إذا كانت العلاقة طردية تامة، ويكون - ١ إذا كانت العلاقة الخطية عكسية سالبة ، ويساوى صفراً إذا كانت معدومة . فهو بذلك ذو خصائص مماثلة لخصائص جيب تمام (جتا) الزاوية الواقعة بين الخط المستقيم والمحور السينى (الزاوية ل) فى الشكل أدناه.



شكل (٢) : جتا ل = ر

لايتأثر الارتباط بتعديل مقياس الرسم (وحدة القياس)، أو نقطة الأصل (الجمع والطرح)، وليست له وحدة قياس . لذلك أصبح معامل الارتباط الخطى هو المقياس الكمى والنوعى للعلاقة الخطية بين المتغيرين. فعدمه، أو ضعفه يعنى عدم ، أو ضعف العلاقة الخطية ، ولكنه ليس دليلاً على عدم وجود أى علاقة ؛ لأن العلاقة قد تكون غير خطية كما أن وجوده لايعنى السببية .

مثال (٤ ، ١٠) :

استخدم مثال (٣) السابق لإيجاد الارتباط بين المتغيرين .

الحل :

من مثال (٣) :

$ع_{س ص} = ٤٠$

$ع^{٢}_{س} = ٨٠$

$ع^{٢}_{ص} = ٢٠$

$$ر = \frac{ع_{س ص}}{ع_{س} ع_{ص}} \qquad (٥)$$

$$= \frac{٤٠}{\sqrt{٨٠ \times ٢٠}}$$

$= ١$ أى أن العلاقة طردية تامة

```
10  REM  PROGRAM TO CALCULATE THE COVARIANCE
30  REM          حساب التغاير بين متغيرين
40  X1=0
50  Y1=0
60  S=0
70  READ N    REM  NO OF OBSERVATIONS
80  FOR I=1 TO N
90   READ X(I),Y(I)  REM المتغيرين
100 NEXT I
110 FOR I=1 TO N
120  X1=X1+X(I)      REM SUM OF X
130  Y1=Y1+Y(I)      REM SUM OF Y
140  S=S+X(I)*Y(I)
150 NEXT I
160 C=(S-(X1*Y1)/N)/(N-1)
170 PRINT 'العمرXالوزن';'العمر';'الوزن';'الرقم'
180 PRINT '____________';'_____';'_____';'_____'
190 FOR I=1 TO N
200  PRINT X(I)*Y(I),Y(I),X(I),I
210 NEXT I
220 PRINT '_____','______','______'
230 PRINT S,X1,Y1
240 PRINT ,C;' = التغاير'
250 DATA 5,20,4,26,7,34,11,38,13,42,15
260 END
```

المخرجات

الرقم	الوزن	العمر	الوزنX العمر
1	20	4	80
2	26	7	182
3	34	11	374
4	38	13	494
5	42	15	630
	50	160	1760

التغاير = 40

البرنامج التالى يقوم بحساب الارتباط الخطى بين متغيرين للبيانات الواردة فى المثال (٤ , ١٠) باستخدام معادلة الارتباط :

$$L = \frac{A}{\sqrt{BC}}$$

حيث :

$$A = S - \frac{X_1 Y_1}{N}$$

$$B = X_2 - \frac{X_1^2}{N}$$

$$C = Y_2 - \frac{Y_1^2}{N}$$

$$S = \sum XY$$

$$X_1 = \sum X$$

$$Y_1 = \sum Y$$

$$X_2 = \sum X^2$$

$$Y_2 = \sum Y^2$$

N = حجم العينة

```
10  REM  PROGRAM TO COMPUTE LINEAR CORRELATION
30  REM  حساب الارتباط الخطى
40  X1=0
50  Y1=0
60  X2=0
70  Y2=0
80  S=0
90  READ N   REM  NO OF OBSERVATIONS
100 FOR I=1 TO N
110  READ X(I),Y(I)
120 NEXT I
124 FOR I=1 TO N
125  X1=X1+X(I)     REM SUM OF X
130  Y1=Y1+Y(I)     REM SUM OF Y
135  X2=X2+X(I)*X(I)
138  Y2=Y2+Y(I)*Y(I)
140  S=S+X(I)*Y(I)
150 NEXT I
160 A=S-X1*Y1/N
170 B=X2-X1*X1/N
180 C=Y2-Y1*Y1/N
190 L=A/SQR(B*C)
195 PRINT 'ص ر٢','س ر٢','س ر ص ر','ص ر','س ر'
198 PRINT '_____','_____','_____','_____','_____'
200 FOR I=1 TO N
210  PRINT Y(I)*Y(I),X(I)*X(I),X(I)*Y(I),Y(I),X(I)
220 NEXT I
230 PRINT '_____','_____','_____','_____','_____'
240 PRINT Y2,X2,S,Y1,X1
250 PRINT ,L;' = الارتباط'
260 DATA 5,20,4,26,7,34,11,38,13,42,15
270 END
```

المخرجات

س ر	ص ر	س ر ص ر	س ر٢	ص ر٢
20	4	80	400	16
26	7	182	676	49
34	11	374	1156	121
38	13	494	1444	169
42	15	630	1764	225
160	50	1760	5440	580

الارتباط = 1

٣ - معنوية الارتباط :

يعتبر معامل ارتباط بيرسون مقدراً للارتباط النظرى (ز) الخاص بالمجتمع الثنائى، الذى سحبت منه العينة ذات الحجم (ن)، بيد أن إعادة السحب قد تؤدى إلى معامل ارتباط آخر. وإذا تم تكرار عملية الاختيار العشوائى للعينات الثنائية، فسوف تكون هناك عدة ارتباطات هى $ر_1$ ، $ر_2$ ، $ر_3$ ، . . . . . وكل منها يمثل تقديراً للارتباط الحقيقى الخاص بالمجتمع (ز).

بيد أن توزيع الارتباطات العينية $ر_1$ ، $ر_2$ ، $ر_3$ ، . . . . . لايتبع لأى توزيع من التوزيعات الإحصائية إذا كانت ز ≠ صفراً. إلا أنه، وبافتراض أن ز = صفراً، وباعتبار أن التوزيع الخاص بالمجتمع ، يخضع للتوزيع الطبيعى فإن :

$$ت = \frac{ر}{\sqrt{\frac{1 - ر^2}{ن - 2}}} \quad (7)$$

تتبع توزيع ت على (ن - ٢) درجات حرية، لذلك تستخدم الإحصائية السابقة لاختبار الفرضية القائلة بأن :

$ف_0$: ز = صفراً

مقابل أى فرضية من البدائل التالية :

$ف_1$: ز ≠ صفر

أو :

$ف_1$: ز > صفر أو $ف_1$: ز < صفر

مثال (١٠,٥) :

اختيرت عينة عشوائية قوامها ٢٧ من بيانات ذات بعدين، فاتضح أن الارتباط الخطى يساوى ٠,٨ فهل يعتبر ذلك دليلاً على وجود ارتباط بمستوى معنوية ٥٪؟

الحل :

$ف_0$: ز = صفراً

$ف_1$: ز ≠ صفراً

إحصائية الاختبار من المعادلة (٧) هى :

$$ت = \frac{ر}{\sqrt{\frac{1 - ر^2}{ن - 2}}} \quad (7)$$

$$= \frac{0,8}{\sqrt{\frac{1 - 0,64}{25}}}$$

$$= \frac{٥ \times ٠,٨}{٠,٦}$$

$$= ٦,٦٧$$

القيمة الحرجة من جدول ت بالملحق رقم (٢) فى نهاية الكتاب على ٢٥ درجات حرية تساوى ٢,٠٦٠ .

وبما أن إحصائية الاختبار أكبر من القيمة الحرجة فلابد من قبول الفرضية البديلة بعد رفض فرضية العدم .

هذا، ويستخدم جدول (١) فى نهاية الفصل لاختبار الفرضيات بطريقة مباشرة .

مثال (٦,١٠) :

اختيرت عينة عشوائية حجمها ٢٨ لبيانات ذات بعدين، فاتضح أن الارتباط الخطى يساوى ٠,٢٣ ، اختبر الفرضية القائلة بأن الارتباط الحقيقى للمجتمع ٠,٣ ، وأوجد حدود الثقة لمعامل ارتباط المجتمع، وذلك بمستوى معنوية ٥٪ .

الحل :

$$أ = ٥٪$$

$$\frac{أ}{٢} = ٢,٥٪$$

باستخدام جدول التوزيع الطبيعى بالملحق (١) نجد أن القيمة الحرجة هى :

$$ى_{٠,٩٧٥} = ١,٩٦$$

أما إحصائية الاختبار فهى :

$$ى = \frac{س - و}{م} \qquad (١١)$$

$$س = \frac{١}{٢} \text{ لن } \frac{١ + ر}{١ - ر} \qquad (٨)$$

ومن جدول (٢) فى نهاية هذا الفصل يتضح أن

$$\text{س} = \frac{\text{١}}{\text{٢}}\ \text{لن}\ \frac{\text{١} + \text{٠,٢٣}}{\text{١} - \text{٠,٢٣}}$$

$$= \text{٠,٢٣٤}$$

وباستخدام نفس الجدول :

$$\text{و} = \frac{\text{١}}{\text{٢}}\ \text{لن}\ \frac{\text{١} + \text{٠,٣}}{\text{١} - \text{٠,٣}}$$

$$= \text{٠,٣١٠}$$

$$\text{م} = \frac{\text{١}}{\sqrt{\text{ن} - \text{٣}}} \qquad (\text{١٠})$$

$$= \frac{\text{١}}{\sqrt{\text{٢٨} - \text{٣}}}$$

$$= \text{٠,٢}$$

وعليه تكون إحصائية الاختبار بعد التعويض فى المعادلة (١١) هى :

$$\text{ى} = \frac{\text{٠,٢٣٤} - \text{٠,٣١٠}}{\text{٠,٢}}$$

$$= -\ \text{٠,٣٨}$$

وبما أن :

$$-\ \text{٠,٣٨} > -\ \text{١,٩٦}$$

فهذا يعنى أنه لا بد من قبول فرضية العدم القائلة بأن ز = ٠,٣ ، بمعنى أنه لايوجد فرق جوهرى بين ٠,٣ والقيمة العينية للارتباط التى تساوى ٠,٢٣

أما حدود الثقة فهى :

$$\text{س} \pm \text{ى}_{\text{٠,٩٧٥}} \times \text{م} \qquad (\text{١٢})$$

٠,٢٣٤ ± ١,٩٦ × ٠,٢

٠,٢٣٤ ± ٠,٣٩٢

وعليه تكون :

- ٠,١٥٨ ≤ و ≤ ٠,٦٢٦

وباستخدام جدول (٢) مرة أخرى بطريقة معاكسة لإيجاد ز لكل حالة باعتبار أن

$$- ٠,١٥٨ = لن \frac{١ + ز}{١ - ز}$$

وأيضاً :

$$٠,٦٢٦ = لن \frac{١ + ز}{١ - ز}$$

يلاحظ أن :

- ٠,١٥٨ تناظرها - ٠,١٥٩

٠,٦٢٦ تناظرها ٠,٥٥٥

وعليه تكون :

- ٠,١٥٩ ≤ ز ≤ ٠,٥٥٥

٤ - اختبار الفرق بين ارتباطين لعينتين:

يجب ألا يكـون الفرق بين ارتباطين لعينتين من نفس المجتمع جوهرياً، ولاختبار ذلك الفرق يستخدم اختبار الفرق بين ارتباطين لعينتين؛ بهدف التحقق من تبعيتهما لمجتمع واحد، أو مجتمعين مختلفين. والمثال التالي يوضح ذلك.

مثال (٧,١٠):

أجريت دراسة لعينتين من طلاب مدرستين لمعرفة الارتباط بين درجات اللغة العربية واللغة الإنجليزية، فكانت النتائج كالآتى :

اسم المدرسة	حجم العينة (ن)	الارتباط (ر) بين اللغتين
مدرسة قيس	٥	٠,٨٩
مدرسة زهير	٨	٠,٦٣

فهل هناك فرق جوهرى بين الارتباطين بمستوى معنوية ٥٪؟

الحل :

إذا كانت :

$ر_{١}$ = ارتباط العينة الأولى.

$ر_{٢}$ = ارتباط العينة الثانية.

$ن_{١}$ = حجم العينة الأولى.

$ن_{٢}$ = حجم العينة الثانية.

ففرضية العدم هى :

$ف_{.} : ر_{١} = ر_{٢}$

والفرضية البديلة هى :

$ف_{١} : ر_{١} \neq ر_{٢}$

$$\frac{١}{ن_{٢} - ٣} = \frac{١}{٥}$$

$$= ٠,٢$$

$$\frac{١}{ن_{١} - ٣} + \frac{١}{ن_{٢} - ٣} = ٠,٥ + ٠,٢$$

$$= ٠,٧$$

إحصائية الاختبار :

$$ى = \frac{١,٤٢٢ - ٠,٧٤١}{\sqrt{٧}}$$

$$= \frac{٠,٦٨١}{\sqrt{٧}}$$

$$= ٠,٢٥٧$$

وبما أن :

$$١,٩٦ > ٠,٢٥٧$$

فليس هناك دليل كافٍ لرفض فرضية العدم، القائلة بأنه لا يوجد فرق جوهرى بين الارتباطين.

٥ ـ معامل ارتباط الرتب لمتغيرين تسلسليين

(Spearman's Rank correlation)

يرتب المتغـيران أولاً تصاعدياً أو تنازلياً، بنفس الطريقة الواردة فى الفصل السابق، ثم يستخرج الفرق (انحراف كل زوج من الأزواج المرتبة) ، ويرمز له بالرمز (ل). وبافتراض أن عدد الفروقات يساوى ن فمعامل سبيرمان لارتباط الرتب هو :

$$ر = \frac{١ - ٦ \sum ل^{٢}}{ن (ن^{٢} - ١)} \qquad (١٦)$$

هذا، ويمتاز معامل ارتباط سبيرمان للرتب بنفس خصائص معامل بيرسون، وتستخدم نفس الأساليب السابقة لاختبارات الفرضيات وفترات الثقة.

مثال (٨,١٠) :

البيانات التالية تمثل درجات عينة قوامها ١٠ أشخاص تقدموا للالتحاق بوظيفة، فأجريت لهم مقابلات شخصية من لجنة مكونة من عضوين، يقوم كل عضو برصد درجة لكل شخص من ١٠ • أوجد معامل سبيرمان لدرجات عضوى اللجنة.

الرقم	درجات العضو الأول	درجات العضو الثانى
١	٩	١٠
٢	٥	٧
٣	٦	٥
٤	١	١
٥	٣	٤
٦	٢	٢
٧	٨	٩
٨	٧	٦
٩	١٠	٨
١٠	٤	٣

الحل :

الرقم	درجات الأول	درجات الثانى	الفرق (ل) الأول - الثانى	مربع الفرق ($ل^2$)
١	٩	١٠	-١	١
٢	٥	٧	-٢	٤
٣	٦	٥	+١	١
٤	١	١	٠	٠
٥	٣	٤	-١	١
٦	٢	٢	٠	٠
٧	٨	٩	-١	١
٨	٧	٦	+١	١
٩	١٠	٨	+٢	٤
١٠	٤	٣	+١	١
المجموع			صفر	١٤

$$\text{ر} = ١ - \frac{٦ \sum \text{ل}^{٢}}{\text{ن} (\text{ن}^{٢} - ١)} \qquad (١٦)$$

$$= ١ - \frac{٦ \times ١٤}{١٠ \times ٩٩}$$

$$= ٠,٩١٥$$

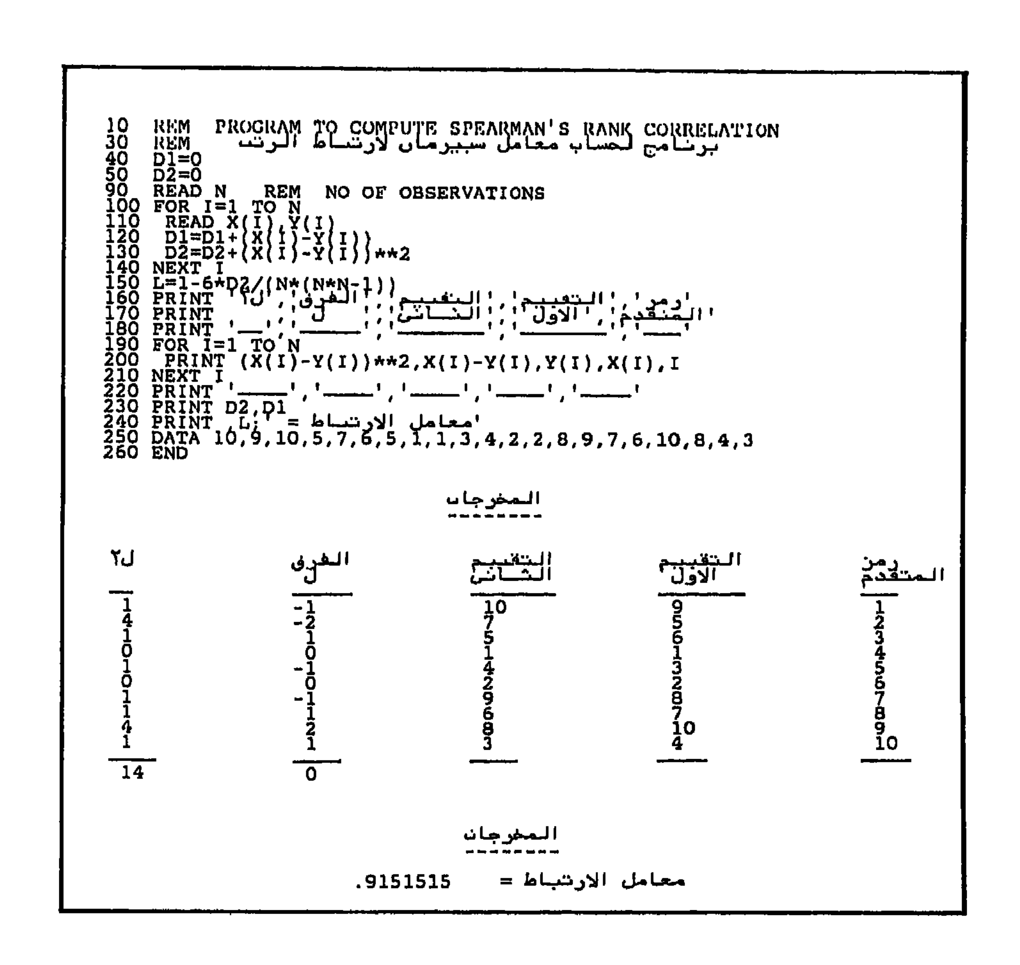
البرنامج التالى يقوم بحساب معامل سبيرمان لارتباط الرتب والبيانات المستخدمة هناهى البيانات بالمثال (٨ , ١٠) السابق.باستخدام معادلة الارتباط للرتب :

$$L = 1 - \frac{6D'^2}{N(N'^2 - 1)}$$

حيث :

D'^2 = مجموع مربعات فروقات الرتب

N = حجم العينة



```
10  REM  PROGRAM TO COMPUTE SPEARMAN'S RANK CORRELATION
30  REM   برنامج لحساب معامل سبيرمان لارتباط الرتب
40  D1=0
50  D2=0
90  READ N    REM  NO OF OBSERVATIONS
100 FOR I=1 TO N
110  READ X(I),Y(I)
120  D1=D1+(X(I)-Y(I))
130  D2=D2+(X(I)-Y(I))**2
140 NEXT I
150 L=1-6*D2/(N*(N*N-1))
160 PRINT 'ل٢','الفرق','التقييم','التقييم','رمز'
170 PRINT ,'ل','الثانى','الاول','المتقدم'
180 PRINT '__','______','______','______','__'
190 FOR I=1 TO N
200  PRINT (X(I)-Y(I))**2,X(I)-Y(I),Y(I),X(I),I
210 NEXT I
220 PRINT '____','____','____','____','____'
230 PRINT D2,D1
240 PRINT L;' = معامل الارتباط'
250 DATA 10,9,10,5,7,6,5,1,1,3,4,2,2,8,9,7,6,10,8,4,3
260 END
```

المخرجات

ل٢	الفرق ل	التقييم الثانى	التقييم الاول	رمز المتقدم
1	-1	10	9	1
4	-2	7	5	2
1	1	5	6	3
0	0	1	1	4
1	-1	4	3	5
0	0	2	2	6
1	-1	9	8	7
1	1	6	7	8
4	2	8	10	9
1	1	3	4	10
14	0			

المخرجات

معامل الارتباط = .9151515

٦ ـ الارتباط الجزئى (PARTIAL CORRELATION):

يستخدم الارتباط الجزئى لقياس العلاقة الخطية بين متغيرين فقط، باعتبار أن بقية المتغيرات ثابتة . لذلك يلجأ الباحثون لاستخدام الارتباط الجزئى ، عندما يكون عدد المتغيرات أكثر من اثنين ، أى أن البيانات ذات عدة أبعاد .

افرض أن هناك ثلاثة متغيرات ، ارتباطاتها فيما بينها على النحو الآتى :

$\text{ر}_{\text{٢١}}$ = الارتباط بين س و ص ($\text{ر}_{\text{س ص}}$)

$\text{ر}_{\text{٣١}}$ = الارتباط بين س و ل ($\text{ر}_{\text{س ل}}$)

$\text{ر}_{\text{٣٢}}$ = الارتباط بين ص و ل ($\text{ر}_{\text{ص ل}}$)

فالارتباط الجزئى بين س و ص باعتبار أن ل ثابت، يرمز له بالرمز $\text{ر}_{\text{٢١و٣}}$ ، أى أن المتغير الذى يكون على يسار الفاصلة هو الثابت . ولقد جاء مفهوم الارتباط الجزئى من أنه الارتباط الخطى بين الانحرافين عن المتغير الثابت ، وبناء على ذلك أصبح فى الإمكان تمثيل الارتباط الجزئى على النحو التالى :

$$\text{ر}_{\text{٢١و٣}} = \frac{\text{ر}_{\text{٢١}} - \text{ر}_{\text{٣١}}\,\text{ر}_{\text{٣٢}}}{\left[(\text{١} - \text{ر}^{\text{٢}}_{\text{٣١}})(\text{١} - \text{ر}^{\text{٢}}_{\text{٣٢}})\right]^{\text{١}/\text{٢}}} \quad (\text{١٧})$$

وبالمثل فإن :

$$\text{ر}_{\text{٣١و٢}} = \frac{\text{ر}_{\text{٣١}} - \text{ر}_{\text{٢١}}\,\text{ر}_{\text{٢٣}}}{\left[(\text{١} - \text{ر}^{\text{٢}}_{\text{٢١}})(\text{١} - \text{ر}^{\text{٢}}_{\text{٣٢}})\right]^{\text{١}/\text{٢}}} \quad (\text{١٨})$$

أما إذا كان عدد المتغيرات أربعة فإن :

$$\text{ر}_{\text{٢١و٣٤}} = \frac{\text{ر}_{\text{٢١و٣}} - \text{ر}_{\text{٤١و٣}}\,\text{ر}_{\text{٤٢و٣}}}{\left[(\text{١} - \text{ر}^{\text{٢}}_{\text{٤١و٣}})(\text{١} - \text{ر}^{\text{٢}}_{\text{٤٢و٣}})\right]^{\text{١}/\text{٢}}} \quad (\text{١٩})$$

أو أن :

$$\text{ر}_{\text{٢١و٣٤}} = \frac{\text{ر}_{\text{٢١و٤}} - \text{ر}_{\text{٣١و٤}}\,\text{ر}_{\text{٣٢و٤}}}{\left[(\text{١} - \text{ر}^{\text{٢}}_{\text{٣١و٤}})(\text{١} - \text{ر}^{\text{٢}}_{\text{٣٢و٤}})\right]^{\text{١}/\text{٢}}} \quad (\text{٢٠})$$

علماً بأن نتيجتى (١٩) و (٢٠) متساويتان .

مثال (٩ ,١٠) :

إذا كانت :

$ر_{٢١} = ٠,٦٣$

$ر_{٣١} = ٠,٨١$

$ر_{٣٢} = ٠,٧٢$

فأوجد $ر_{٢١و٣}$

الحل :

$$ر_{٢١و٣} = \frac{ر_{٢١} - ر_{٣١} ر_{٣٢}}{[(١ - ر_{٣١}^{٢})(١ - ر_{٣٢}^{٢})]^{\frac{١}{٢}}} \qquad (١٧)$$

وبالتعويض فى المعادلة السابقة :

$$ر_{٢١,٣} = \frac{٠,٦٣ - ٠,٨١ \times ٠,٧٢}{[(١ - (٠,٨١)^{٢}) (١ - (٠,٧٢)^{٢})]}$$

$$= \frac{٠,٠٤٦٨}{٠,٤٠٧}$$

$$= ٠,١١٥$$

هذا، وتستخدم جداول اختبار الارتباط الخطى لاختبار معنوية اختلاف الارتباط الجزئى عن الصفر، وذلك باعتبار أن حجم العينة ن - ١ إذا كان عدد المتغيرات ثلاثة، و ن - ٢ إذا كانت عدد المتغيرات أربعة.

٧ - الارتباط الثنائي التسلسل (Biserial correlation) :

إذا كان المتغير س $_{ر}$ متصلاً بينما كان المتغير ص $_{ر}$ ثنائي التسلسل، كأن تقسم المدينة إلى منطقتين، أو حالات الإجابة بنعم أو لا، أو الحالة الاجتماعية (متزوج ، أو غير متزوج)، وإذا كانت :

سَ $_{١}$ = الوسط الحسابي للمتغير المتصل عندما كانت ص $_{ر}$ = ص $_{١}$

سَ $_{٢}$ = الوسط الحسابي المقابل لصفة التقسيم الثانية (ص $_{٢}$)

$$ح = \frac{ص_{١}}{ص_{١} + ص_{٢}} \qquad (٢١)$$

$$حَ = \frac{ص_{٢}}{ص_{١} + ص_{٢}}$$

$$= ١ - ح \qquad (٢٢)$$

فالارتباط الخطي بين المتغير النسبي المتصل (س $_{ر}$) والمتغير الخاص بصفتي التقسيم (ص $_{ر}$) هو :

$$ر = \left(\frac{سَ_{١} - سَ_{٢}}{ع_{س}} \right) \times (ح حَ)^{\frac{١}{٢}} \qquad (٢٣)$$

هذا، ويلاحظ أن المتغير الثنائي التسلسل لايدخل في العمليات الحسابية الخاصة باستخراج الارتباط، أما اختبارات المعنوية وحدود الثقة فتستخدم لها نفس المعادلات الخاصة بمعامل بيرسون للارتباط الخطي.

مثال (١٠ , ١٠) :

البيانات التالية تمثل عدد القتلى في حوادث المرور، موزعين حسب مكان الوفاة خلال الفترة من أبريل حتى ديسمبر ١٩٨٠م.

والبيانات[1] هي :

الشهر	عدد القتلى قبل الوصول للمستشفى	عدد القتلى بعد الوصول للمستشفى
أبريل	٢١	١٨
مايو	٢٦	٢١
يونيو	٣٦	٢٦
يوليو	٤٦	٣٣
أغسطس	٣٩	٢٩
سبتمبر	٣٤	٢٤
أكتوبر	٣٠	١٩
نوفمبر	١٩	١٤
ديسمبر	١٩	١٥
المجموع	٢٧٠	١٩٩

أوجد الارتباط بين عدد القتلى ومكان الوفاة .

الحل :

$$ر = \frac{\bar{س}_١ - \bar{س}_٢}{ع_س} \times \sqrt{ح\,خ} \qquad (٢٣)$$

$$ع_س = ٨,٨٦٨$$

$$\bar{س}_١ = \frac{٢٧٠}{٩}$$

$$= ٣٠$$

$$\bar{س}_٢ = \frac{١٩٩}{٩}$$

$$= ٢٢,١١$$

$$ح = \frac{ص_١}{ص_١ + ص_٢} \qquad (٢١)$$

(١) المصدر : الإدارة العامة للمرور بالرياض : بحث «حوادث السيارات والأضرار الصحية الناتجة عنها». الرياض ، ١٩٨١، صفحة (٢١).

$$= \frac{٢٧٠}{٢٧٠ + ١٩٩}$$

$$= ٠,٥٧٦$$

$$ح = \frac{ص_٢}{ص_١ + ص_٢} \qquad (٢٣)$$

$$= \frac{١٩٩}{٤٦٩}$$

$$= ٠,٤٢٤$$

$$∴ ر = \frac{٣٠ - ٢٢,١١}{٨,٨٦٨} \times \sqrt{٠,٥٧٦ \times ٠,٤٢٤}$$

$$= ٠,٤٤$$

هذا، ويلاحظ من جدول (١) فى نهاية هذا الفصل أن القيمة الحرجة لاختبار الفرضية :

$ف_٠$: ز = صفراً

مع الفرضية البديلة :

$ف_١$: ز > صفر

وبمستوى ٥٪ تساوى ٠,٤٩٧ ، وفى ذلك دلالة على خطورة الإصابات التى يصعب إسعافها فى أكثر الحالات.

البرنامج التالى يقوم بحساب الارتباط الثنائى مستخدماً البيانات الواردة بالمثال (١٠ , ١٠) السابق وباستخدام المعادلة :

$$D = \frac{B - C}{V} \sqrt{AE}$$

حيث :

$$B = \frac{X_1}{N}$$

$$C = \frac{Y_1}{N}$$

$$A = \frac{X_1}{X_1 + Y_1}$$

$$E = 1 - A$$

$$V = \sqrt{\left(X_2 + Y_2 - (X_1 + Y_1)/N\right) / (N-1)}$$

$X_1 \; Y_1 =$ مجاميع

$X_2 \; Y_2 =$ مجاميع مربعات

لابد من حساب الانحراف المعيارى (V) هنا الذى يساوى 8.868 المستخدم فى السطر رقم 220 وهو عبارة عن :

$$\left[(21)^2 + (26)^2 + (36)^2 + \ldots + (19)^2 + (18)^2 + (21)^2 + (15)^2 - \Sigma \; (21 + 26 + 36 + \ldots + 15)^2 / N \right]$$

$$\div \; (N-1)$$

حيث :

$N = 18$

مستخدماً V الواردة فى المعادلة السابقة .

```
10  REM  PROGRAM TO COMPUTE BISERIAL CORRELATION
30  REM    برنامج لحساب الارتباط الثنائى
40  X1=0
50  Y1=0
60  READ N   REM  NO OF OBSERVATIONS
70 PRINT 'العدد ٢','العدد ١','الشهر'
80 PRINT '________','________','________'
90 FOR I=1 TO N
100  READ X,Y
110  PRINT Y,X,I
120  X1=X1+X
130  Y1=Y1+Y
140 NEXT I
150 A=X1/(X1+Y1)
190 B=X1/N
200 C=Y1/N
210 E=1-A
220 D=(B-C)/8.868*SQR(A*E)
230 PRINT '________','________'
240 PRINT Y1,X1,'المجموع'
250 PRINT ,D;' = الارتباط الثنائى'
260 DATA 9,21,18,26,21,36,26,46,33,39,29,34,24,30,19,19,14,19,15
270 END
```

المخرجات

الشهر	العدد ١	العدد ٢
1	21	18
2	26	21
3	36	26
4	46	33
5	39	29
6	34	24
7	30	19
8	19	14
9	19	15
المجموع	270	199

الارتباط الثنائى = .4396694

أما إذا كان المتغير الثنائى متصلاً فى الأصل، ولكنه قسم اصطناعياً لصفتين، كأن يقسم الدخل إلى قسمين، أحدهما أكثر من ١٠٠٠٠ ريال شهرياً، والثانى أقل من ذلك المقدار، أو عند تقسيم الأعمار لأكثر من ٢٠ سنة وأقل من ذلك ـ فالارتباط الثنائى التسلسل هو :

$$\text{ر} = \frac{\text{ح ح}'}{\text{ى}'} \cdot \frac{(\bar{\text{س}}_{1} - \bar{\text{س}}_{2})}{\text{ع س}} \qquad (٢٤)$$

حيث :

$$\text{ى}' = \frac{1}{\sqrt{2\,\text{ط}}}\ \text{هـ}^{\left(-\frac{\text{ى}^{2}}{2}\right)} \qquad (٢٥)$$

علماً بأن :

$$\text{ط} = \frac{22}{7}$$

$$\text{هـ} = 2{,}71828$$

أمـا قيمة ى فتستخرج من جدول التوزيع الطبيعى بالملحق؛ لأنها تساوى القيمة التى ينقسم عندها المنحنى الخاص بالتوزيع الطبيعى إلى جزأين متكاملين هما ح وحَ.

مثال (١١,١٠) :

توضح البيانات [٢] التالية عدد المصابين والقتلى فى حوادث المرور، حسب وقت وقوع الحادث، خلال الفترة من أبريل حتى ديسمبر ١٩٨٠.

والبيانات هى :

وقت وقوع الحادث	عدد المصابين والقتلى
٦ - ٨ صباحاً	٧٦
٨ - ١٠ صباحاً	٧٠
١٠ - ١٢ ظهراً	٢٨
١٢ - ٢ ظهراً	٢٨
٢ - ٤ عصراً	٦٨
٤ - ٦ مساء	٧٤
٦ - ٨ مساء	٥٤
٨ - ١٠ مساء	٢٨
١٠ - ١٢ صباحاً	٦
١٢ - ٢ صباحاً	٢
٢ - ٤ صباحاً	٢
٤ - ٦ صباحاً	١٠
المجموع	٤٤٦

أوجد الارتباط بين عدد المصابين والقتلى من جهة ووقت وقوع الحادث حسب النهار والليل (أى قبل وبعد السادسة مساء) من جهة أخرى.

الحل :

الوقت متصل، إلا أنه قسم اصطناعياً ألى جزأين، هما : قبل، وبعد الغروب، وعليه تستخدم المعادلة

$$ر_{.} = \frac{ح\, حَ}{ىَ} \cdot \frac{(\bar{س}_{١} - \bar{س}_{٢})}{ع_{س}} \qquad (٢٤)$$

(٢) المصدر : نفس المصدر السابق صفحة (٣٧).

$\bar{س}_{١}$ = الوسط الحسابى لحوادث النهار.
$\bar{س}_{٢}$ = الوسط الحسابى لحوادث الليل.
ح = نسبة حوادث النهار للعدد الكلى.
حَ = ١ - ح وهى نسبة حوادث الليل للعدد الكلى.

إذاً :

$$\bar{س}_{١} = \frac{٧٤ + ٦٨ + ٢٨ + ٢٨ + ٧٠ + ٧٦}{٦}$$

= ٥٧,٣٣٣

$$\bar{س}_{٢} = \frac{١٠ + ٢ + ٢ + ٦ + ٢٨ + ٥٤}{٦}$$

= ١٧

$$ح = \frac{٣٤٤}{٤٤٦}$$

= ٧٧١و٠

حَ = ١ - ح

= ٠,٢٢٩

وباستخدام جدول التوزيع الطبيعى بالملحق رقم (١)، يلاحظ أن القيمة التى تناظر ٠,٧٧١ هى :

ى = ٠,٧٤

وعليه :

$$ىَ = \frac{١}{\sqrt{٢ ط}} هـ^{-\frac{ى^{٢}}{٢}} \qquad (٢٥)$$

$$= \frac{١}{\sqrt{٢ \times \frac{٢٢}{٧}}} هـ^{-\frac{(٠,٧٤)^{٢}}{٢}}$$

$= ٠,٣٠٣$

$ع_{س} = ٥٩,٥٦$

$$ر = \frac{(١٧ - ٥٧,٣٣٣) \times ٠,٢٢٩ \times ٠,٧٧١}{٥٩,٥٦ \times ٠,٣٠٣}$$

$= ٠,٧٩٥$

هذا ، وتجدر الاشارة إلى أن الارتباط الثنائى التسلسل قد يزيد على الواحد إذا كان توزيع المتغير الثنائى التسلسل بعيداً جداً عن التوزيع الطبيعى. كذلك لايمكن استخدام الاختبارات السابقة فى هذا النوع من الارتباطات الخطية.

٨ ـ معامل الارتباط الرباعى للتقسيم الاصطناعى

(Tetrachoric Correlation)

إذا قسم المتغيران المتصلان تقسيماً اصطناعياً، بحيث ينقسم المتغير الأول إلى قسمين هما $س_١$ و $س_٢$ وينقسم المتغير الثانى إلى قسمين أيضاً هما $ص_١$ و $ص_٢$، فسوف تنقسم البيانات إلى أربع خلايا على النحو التالى :

ص / س	$ص_١$	$ص_٢$
$س_١$	ب	جـ
$س_٢$	د	ل

ويكون الارتباط الخطى بين المتغيرين على النحو الآتى :

$$ر = جتا \left[\frac{ط}{١ + \left(\frac{ب\, ل}{جـ\, د}\right)^{\frac{١}{٢}}} \right] \qquad (٢٦)$$

مثال (١٢، ١٠) :

البيانات التالية عبارة عن مائة شخص، تم تقسيمهم إلى مجموعات حسب مستوى الدخل والعمر. أوجد الارتباط بين الدخل والعمر، والبيانات هى :

الدخل \ العمر	أقل من ٤٠ سنة	٤٠ سنة فما فوق
أقل من ١٠٠٠٠ ريال شهرياً	٤٩	١٥
١٠٠٠٠ ريال شهرياً فأكثر	١٠	٢٦

الحل :

ب = ٤٩

جـ = ١٥

د = ١٠

ل = ٢٦

$$ر = جتا \left[\frac{ط}{1 + \left(\frac{ب\ ل}{جـ\ د} \right)^{\frac{1}{2}}} \right] \qquad (٢٦)$$

$$= جتا \left[\frac{ط}{1 + \left(\frac{49 \times 26}{15 \times 10} \right)^{\frac{1}{2}}} \right]$$

ر = ٠,٩٤٦

٩ - الارتباط بين المتغيرات الاسمية (phi and Cramer's V correlations) :

إذا كانت :

ب، جـ ، د، ل عبارة عن بيانات بخلايا جدول ٢×٢ لمتغيرين اسميين على النحو التالى :

$$\begin{bmatrix} \text{جـ} & \text{ب} \\ \text{ل} & \text{د} \end{bmatrix}$$

فإن معامل فاى للارتباط هو :

$$\text{ر} = \frac{\text{جـ د} - \text{ب ل}}{[(\text{ب} + \text{جـ})\ (\text{د} + \text{ل})\ (\text{ب} + \text{د})\ (\text{جـ} + \text{ل})]^{\frac{1}{2}}} \qquad (27)$$

هذا، ويعتبر معامل فاى أحد أنواع معامل بيرسون.

مثال (١٣ , ١٠) :

تمثل البيانات التالية [٣] عدد المصابين والقتلى لكل مائة حادث مرورى فى مدينتى الرياض وجدة خلال عام ١٤٠١هـ . أوجد الارتباط بين نوع الحادث (إصابة أو قتل)، والمدينة . والبيانات هى :

نوع الحادث / المدينة	إصابات	قتلى
الرياض	٥٠	٤
جدة	١١٨	١٣

الحل :

$$\text{ر} = \frac{\text{جـ د} - \text{ب ل}}{[(\text{ب} + \text{جـ})\ (\text{د} + \text{ل})\ (\text{ب} + \text{د})\ (\text{جـ} + \text{ل})]^{\frac{1}{2}}} \qquad (27)$$

ب = ٥٠

جـ = ٤

د = ١١٨

ل = ١٣

(٣) المصدر : الإدارة العامة للمرور ـ النشرة الإحصائية لعام ١٤٠١هـ ـ الرياض (١٤٠٢هـ) صفحة (١٥).

وبالتعويض فى (٢٧) سابقاً تكون :

$$ر = \frac{٤ \times ١١٨ - ٥٠ \times ١٣}{[(٥٠ + ٤)(١١٨ + ١٣)(٥٠ + ١١٨)(٤ + ١٣)]^{١/٢}}$$

$$= \frac{-١٧٨}{[٥٤ \times ١٣١ \times ١٦٨ \times ١٧]^{١/٢}}$$

$$= \frac{-١٧٨}{٤٤٩٤,٨١}$$

$$= -٠,٠٤$$

هذا، وتعتبر العلاقة جوهرية (أكبر من الصفر) إذا ثبت ذلك باختبار الاستقلال بجداول التوافق؛ وذلك لأن :

$$ر = \sqrt{\frac{ك^٢}{ن}} \qquad (٢٨)$$

حيث $ك^٢$ هى إحصائية الاختبار التى تتبع توزيع مربع كاى ، أما ن فهى حجم العينة هذا، وتتراوح قيمة الارتباط هنا بين + ١ و - ١ .

أمـا إذا كان جدول التـوافق أكـبر من ٢ × ٢ فمن الأفضـل استخـدام معامل التوافق (coefficient of contingency) بدلاً من معامل فاى. ويعرف معامل التوافق بأنه :

$$ر = \sqrt{\frac{ك^٢}{ن + ك^٢}} \qquad (٢٩)$$

هذا، وقد عدل معامل التوافق لتتراوح قيمته بين + ١ و - ١ . ويعرف الارتباط المعدل بمعامل كرامر (cramer V) الذى يكون على النحو التالى :

$$ر = \sqrt{\frac{ك^٢}{ن (ل - ١)}} \qquad (٣٠)$$

حيث ل هو عدد الصفوف أو الأعمدة أيهما أقل .

جدول (١) القيم الحرجة لاختبار*

للاختبار ذى الجانبين، a قيمتها ضعف القيمة المسجلة عند عنوان العمود الذى له قيمة r الحرجة، لذلك لقيمة

a =.05

اختار العمود 025.

n \ α	.05	.025	.005
5	.805	.878	.959
6	.729	.811	.917
7	.669	.754	.875
8	.621	.707	.834
9	.582	.666	.798
10	.549	.632	.765
11	.521	.602	.735
12	.497	.576	.708
13	.476	.553	.684
14	.457	.532	.661
15	.441	.514	.641
16	.426	.497	.623
17	.412	.482	.606
18	.400	.468	.590
19	.389	.456	.575
20	.378	.444	.561
25	.337	.396	.505
30	.306	.361	.463
35	.283	.334	.430
40	.264	.312	.402
50	.235	.279	.361
60	.214	.254	.330
80	.185	.220	.286
100	.165	.196	.256

* المصدر : بول ج. هويل: المبادىء الأولية فى الإحصاء، ترجمة د. بدرية عبدالوهاب ود. محمد كامل الشربينى، الطبعة الرابعة، مطابع وايل للكتب العربية - نيويورك - ١٩٨٤، صفحة ٣٢٧.

جدول (٢)

جدول تحويل ر إلى ى (تحويل فشر لمعامل الارتباط)**

r	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
.0	0.000	0.010	0.020	0.030	0.040	0.050	0.060	0.070	0.080	0.090
.1	.100	.110	.121	.131	.141	.151	.161	.172	.182	.192
.2	203	.213	.224	.234	.245	.255	.266	.277	.288	.299
.3	.310	.321	.332	.343	.354	.365	.377	.388	.400	.412
.4	.424	.436	.448	.460	.472	.485	.497	.510	.523	.536
.5	.549	.563	.576	.590	.604	.618	.633	.648	.662	.678
.6	.693	.709	.725	.741	.758	.775	.793	.811	.829	.848
.7	.867	.887	.908	.929	.950	.973	.996	1.020	1.045	1.071
.8	1.099	1.127	1.157	1.188	1.221	1.256	1.293	1.333	1.376	1.422
r	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
.90	1.472	1.478	1.483	1.488	1.494	1.499	1.505	1.510	1.516	1.522
.91	1.528	1.533	1.539	1.545	1.551	1.557	1.564	1.570	1.576	1.583
.92	1.589	1.596	1.602	1.609	1.616	1.623	1.630	1.637	1.644	1.651
.93	1.658	1.666	1.673	1.681	1.689	1.697	1.705	1.713	1.721	1.730
.94	1.738	1.747	1.756	1.764	1.774	1.783	1.792	1.802	1.812	1.822
.95	1.832	1.842	1.853	1.863	1.874	1.886	1.897	1.909	1.921	1.933
.96	1.946	1.959	1.972	1.986	2.000	2.014	2.029	2.044	2.060	2.076
.97	2.092	2.109	2.127	2.146	2.165	2.185	2.205	2.227	2.249	2.273
.98	2.298	2.323	2.351	2.380	2.410	2.443	2.477	2.515	2.555	2.599
.99	2.646	2.700	2.759	2.826	2.903	2.994	3.106	3.250	3.453	3.800

** المصدر: محمد صبحى أبو صالح وعدنان محمد عوض: مقدمة فى الإحصاء، دارجون وايلى وأبنائه - نيويورك، ١٩٨٣ صفحة ٣١٥.

تمارين

١ - البيانات التالية عبارة عن العمر بالسنوات ، وقوة التحمل للعمل الشاق لعشرة عمال تم اختيارهم كعينة عشوائية بأحد المصانع . والبيانات هى :

العمر بالسنوات	الزمن بالدقائق
٣٦	٥
٢٥	٨
٣٩	٤
١٩	٩
٣٠	٥
٢١	١٠
٢٦	٧
٤١	٢
٥٦	٤
٣١	٦

١ - ارسم لوحة التشتت لهذه البيانات.
٢ - ما هو التغاير؟
٣ - ماهى مزايا الارتباط الخطى ، وماعلاقته بالتغاير؟
٤ - حلل لوحة التشتت لبيانات السؤال الأول.
٥ - أوجد التغاير والارتباط الخطى لبيانات السؤال الأول.

٦ - إذا كانت :
$س_ر$ = ١ ، ٢ ، ٣ ، ٤ ، ٥ ، ٦ ، ٧ ، ٨
$ص_ر$ = ٦ ، ١٠ ، ٦ ، ١٠ ، ٦ ، ١٠ ، ٦ ، ١٠
فأوجد الارتباط الخطى بين المتغيرين.

٧ - البيانات التالية تمثل تكلفة الصيانة بآلاف الريالات سنوياً، والعمر لعدد من الناقلات التى اختيرت كعينة عشوائية ، والبيانات هى :

التكلفة	العمر بالسنوات
٧	٣
٤	٥
٤	٤
٩	٧
١	٢
٨	٦

فهل هناك ارتباط بين المتغيرين يختلف عن الصفر بمستوى معنوية ٥٪؟

٨ - هل يختلف ارتباط بيانات السؤال السابق عن ٣٥, . بمستوى معنوية ٥٪؟

٩ - إذا كانت :

س ر = ٩، ١٠، ١١، ١٢، ١٣، ١٤، ١٥، ١٦، ١٧، ١٨، ١٩

ص ر = - ١٠، - ٨، - ٦، -٤، - ٢، ٠، ٢، ٤، ٦، ٨، ١٠

فأوجد الارتباط الخطى بين المتغيرين، واختبر الفرضية القائلة بأنه لا يوجد ارتباط بين المتغيرين.

١٠ - إذا كانت :

س ر = ١٨، ٢٠، ٢٢، ٢٤، ٢٦، ٢٨، ٣٠، ٣٢، ٣٤، ٣٦، ٣٨

ص ر = ٠، ٢، ٤، ٦، ٨، ١٠، ١٢، ١٤، ١٦، ١٨، ٢٠

فأوجد الارتباط الخطى بين المتغيرين، وقارن بينه وبين الارتباط فى السؤال السابق، ووضح أسباب العلاقة بينهما.

١١ - توضح البيانات التالية مبيعات ثلاثة مراكز لتوزيع إحدى المجلات خلال ستة أشهر. اختبر الفرضية القائلة بأن التوزيع لا يتحسن بمرور الزمن لكل مركز بمستوى معنوية ٥٪، والبيانات هى :

المبيعات / الشهر	المركز الأول بآلاف الريالات	المركز الثانى بآلاف الريالات	المركز الثالث بآلاف الريالات
١	١٦	٥	٩
٢	١٤	٦	١١
٣	١١	٧	١٠
٤	١٢	٦	١٢
٥	١٠	٩	١١
٦	١٠	١٠	١٠

١٢ ـ أوجد الارتباط الخطى بين كل مركزين، واختبر فرضية وجود ارتباط بينهما، وفسر النتائج بعد الاختبار.

١٣ ـ استخدم بيانات السؤال السادس لإيجاد معامل سبيرمان للارتباط وقارن بين النتيجتين.

١٤ ـ استخدم بيانات السؤال التاسع لإيجاد معامل سبيرمان للارتباط، وقارن بين النتيجتين.

١٥ ـ استخدم بيانات السؤال العاشر لإيجاد الارتباط بين المتغيرين بمعامل سبيرمان، وقارن بين النتيجتين.

١٦ ـ استخدم بيانات السؤال الأول لإيجاد معامل سبيرمان للارتباط، وقارن بين النتيجتين (مع السؤال الخامس).

١٧ ـ استخدم بيانات السؤال الحادى عشر لإيجاد معامل سبيرمان وقارن بين النتيجتين.

١٨ ـ متى يتساوى معامل بيرسون مع معامل سبيرمان لنفس البيانات ؟

١٩ ـ إذا كانت :ـ

١ تعنى المركز الأول فى السؤال الحادى عشر.

٢ تعنى المركز الثانى فى السؤال الحادى عشر.

٣ تعنى المركز الثالث فى السؤال الحادى عشر

٤ تعنى الزمن فى السؤال الحادى عشر.

فأوجد :

أ ـ ر ٣,٢١

ب ـ ر ٤,٢١

جـ ـ ر ٣,٤١

د ـ ر ٢,٣١

هـ ـ ر ٣,٢٤

و ـ ر ١,٣٤

ز ـ ر ٤٢,٣١

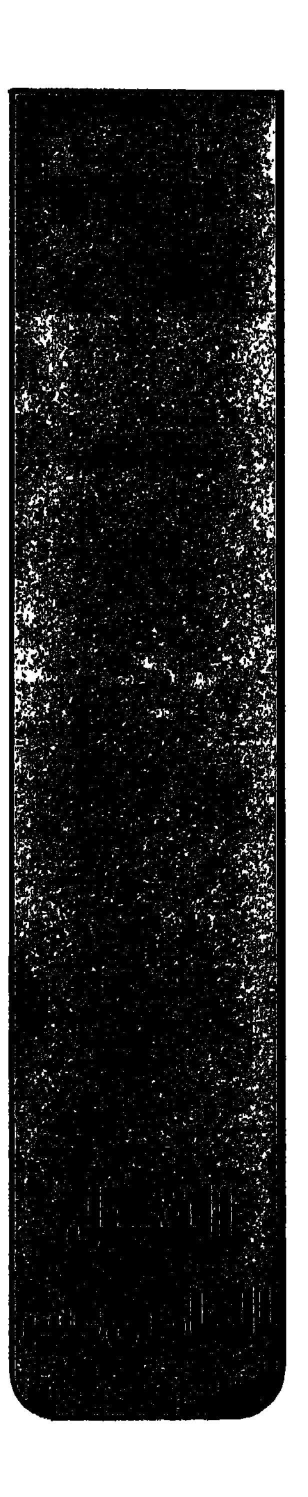
ح ـ ر ٣١,٢٤

ط ـ ر ٤١,٣٢

٢٠ ـ البيانات التالية عبارة عن عينة من المتعلمين، وعينة من الأميين، وعدد المدخنين من كل عينة، فهل هناك ارتباط بين التعليم وعادة التدخين؟ والبيانات هى :

صفة التدخين / صفة التعليم	عدد المدخنين	عدد غير المدخنين
متعلم	١٠	٤٠
أمى	٤٥	٥

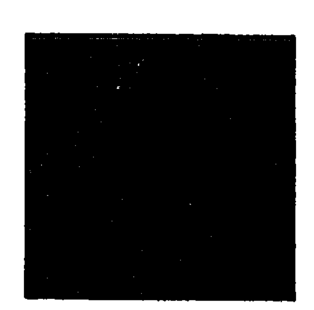
٢١ ـ اكتب برنامجاً بلغة بيسك لرسم بيانى للبيانات الواردة فى السؤال (١).
٢٢ ـ اكتب برنامج بيسك لإيجاد الارتباط الخطى للبيانات بالسؤال (١).
٢٣ ـ استخدم البيانات بالسؤال (٦) فى برنامج لإيجاد التغاير والارتباط الخطى.
٢٤ ـ استخدم البيانات الواردة بالسؤال (٦)، واكتب برنامج بيسك لإيجاد معامل سبيرمان للارتباط.



الانحدار الخطى
(Linear Regression)









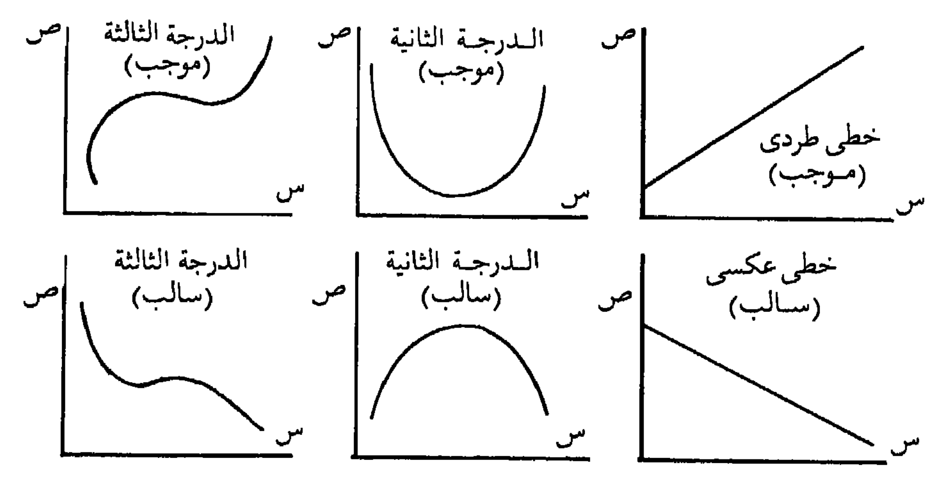
الانحدار الخطى

(Linear Regression)

١ ـ مفهوم الانحدار:

يهدف خط الانحدار الخاص بمتغيرين متصلين إلى تحديد العلاقة بين القيم العينية الثنائية ($س_ر$ ، $ص_ر$)، باعتبار أن $س_ر$ دالة للمتغير $ص_ر$، بمعنى أن المتغير $س_ر$ مسبب (مستقل)، والمتغير $ص_ر$ متغير تابع.

يتلخص دور الانحدار فى ثلاث مهام أساسية، أولاها : الوصف الخاص بظاهرة معينة، وثانيتها : التحكم فى المتغير التابع بواسطة المتغير المستقل، وثالثتها : تقدير (تنبؤ) بعض قيم المتغير التابع بعد تحديد قيم معينة للمتغير المستقل. هذا، ويأتى ذلك بعد تحديد العلاقة بين المتغيرين، والتى تتمثل فى معادلة قد تكون خطية أو غير خطية، اعتماداً على الشكل الذى تبينه لوحه الانتشار. وفيها يلى بعض الأمثلة لنماذج مختلفة :



شكل (١) نماذج لبعض المنحنيات.

هذا، وتنحصر مهمة هذا الفصل فى الانحدار الخطى الذى يكون على النحو التالى :

ص $_{ر}$ = (أ) + (ب) س $_{١ر}$ + (جـ) س $_{٢ر}$ + (د) س $_{٣ر}$ + ٠٠٠٠٠ + خ (١)

حيث :

ص $_{ر}$ متغير تابع

س $_{١ر}$، س $_{٢ر}$، س $_{٣ر}$، ٠٠٠ مجموعة المتغيرات المستقلة.

خ الخطأ العشوائى.

(أ) ، (ب) ، (جـ) ، (د) ٠٠٠٠ هى معالم (ثوابت) المعادلة الواجب تقديرها.

٢ - معادلة الانحدار الخطى البسيط:

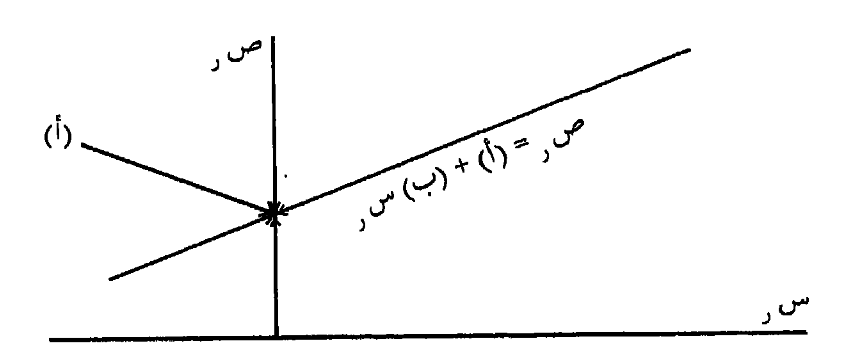
تكون معادلة الانحدار الخطى البسيط على النحو التالى :

ص $_{ر}$ = (أ) + (ب) س $_{ر}$ (٢)

حيث :

(أ) هى الجزء المقطوع من المحور الصادى (انظر الشكل) ، أو هى المعدل العام كما يطلق عليها فى بعض الحالات.

(ب) هى ميل خط الانحدار، أو كمية التغير التى تطرأ على المتغير التابع (ص $_{ر}$) إذا تغير المستقل (س $_{ر}$) بوحدة واحدة.



شكل (٢) : خطة الانحدار البسيط

إذاً تتغير المعادلة بتغير قيمة (أ) ، أو تغير قيمة (ب) : معلمى المعادلة . لذلك فإن المعادلة الخطية المستخرجة من القيم العينية هى معادلة تقدير الثابتين (أ) و(ب)، اعتماداً على تلك القيم . هذا، وقد تختلف التقديرات باختلاف العينات المسحوبة من نفس المجتمع ؛ وعليه فالمعادلة التى تحسب معالمها من القيم العينية هى معادلة تقديرية، وتكون على النحو التالى :

$$ص'_ر = أ + ب\ س_ر \qquad (٣)$$

حيث :

$ص'_ر$ قيمة تقديرية للقيمة الحقيقية $ص_ر$

أ و ب قيمتان تقديريتان للمعلمين (أ) و (ب) على التوالى.

يجب أن تكون $ص'_ر$ أفضل مقدر لنظيرتها $ص_ر$. أى أن الانحراف - أو الخطأ - الذى يرمز إليه بالرمز (خ) بين القيمة الحقيقية للمجتمع والقيمة التقديرية من المعادلة يجب أن يكون فى أدنى حد ممكن .

بيد أن مجموع تلك الأخطاء (م خ) يساوى صفراً، بمعنى أن :

$$م\ خ = \sum (ص_ر - ص'_ر) = صفراً \qquad (٤)$$

إذاً لا بد من اللجوء إلى مجموع مربعات تلك الأخطاء (م م خ) الذى يعرف بأنه :

$$م\ م\ خ = \sum (ص_ر - ص'_ر)^2 \qquad (٥)$$

هذا، وتعرف النظرية التى تعتمد على استخدام أدنى قيمة لمجموع مربعات الأخطاء لتقدير (أ) و (ب) بنظرية المربعات الصغرى (Least Squares Method) . كذلك يتضح من تعويض المعادلة (٣) فى المعادلة (٥) أن :

$$م\ م\ خ = \sum (ص_ر - (أ + ب\ س_ر))^2$$

$$= \sum (ص_ر - أ - ب\ س_ر)^2 \qquad (٦)$$

وبتطبيق التفاضل الجزئى على المعادلة رقم (٦) بالنسبة إلى أ تارة، وبالنسبة إلى ب تارة أخرى مع معادلة ناتج كل تفاضل إلى الصفر بهدف إيجاد أصغر قيمة لمجموع مربعات الأخطاء، يكون الناتج هو المعادلتين الآتيتين التاليتين اللتين يطلق عليهما اسم المعادلتين الطبيعيتين (Normal Equations) :

$$\sum ص_ر = أ ن + ب \sum س_ر \qquad (٧)$$

$$\sum س_ر ص_ر = أ \sum س_ر + ب \sum س_ر^٢ \qquad (٨)$$

وبحل المعادلتين (٧) و (٨) آنياً تكون :

$$أ = \bar{ص} - ب \bar{س} \qquad (٩)$$

$$ب = \frac{\sum س_ر ص_ر - \frac{\sum س_ر \sum ص_ر}{ن}}{\sum س_ر^٢ - \frac{(\sum س_ر)^٢}{ن}}$$

$$= \frac{\sum (س_ر - \bar{س})(ص_ر - \bar{ص})}{\sum (س_ر - \bar{س})^٢} \qquad (١٠)$$

إذاً :

$$ب = \frac{ع_{س ص}}{ع_س^٢} \qquad (١١)$$

مثال (١ , ١١) :

البيانات التالية تمثل أرباح إحدى المؤسسات خلال ٧ سنوات ، وتكلفة الدعاية فى كل سنة من تلك السنوات . أوجد معادلة الانحدار الخطى البسيط للربح على الدعاية ، ثم استخدم تلك المعادلة لتقدير الربح إذا كانت تكلفة الدعاية ١٨ ألف ريال فى سنة ما . والبيانات هى :

الدعاية ($س_ر$) بآلاف الريالات	الربح ($ص_ر$) بآلاف الريالات
١٠	٢٥
١٥	٣٥
١٤	٣٠
١٦	٤٠
١١	٢٤
٢٠	٥٠
١٩	٤٨

الحل :

يتضح من البيانات السابقة أن :

ن = ٧

أما بقية المجاميع فيمكن الحصول عليها على النحو الآتى :

رقم المشاهدة	$\text{س}_{\text{ر}}$	$\text{ص}_{\text{ر}}$	$\text{س}_{\text{ر}}^{٢}$	$\text{ص}_{\text{ر}}^{٢}$	$\text{س}_{\text{ر}}\text{ص}_{\text{ر}}$
١	١٠	٢٥	١٠٠	٦٢٥	٢٥٠
٢	١٥	٣٥	٢٢٥	١٢٢٥	٥٢٥
٣	١٤	٣٠	١٩٦	٩٠٠	٤٢٠
٤	١٦	٤٠	٢٥٦	١٦٠٠	٦٤٠
٥	١١	٢٤	١٢١	٠٥٧٦	٢٦٤
٦	٢٠	٥٠	٤٠٠	٢٥٠٠	١٠٠٠
٧	١٩	٤٨	٣٦١	٢٣٠٤	٩١٢
المجموع	١٠٥	٢٥٢	١٦٥٩	٩٧٣٠	٤٠١١

$$\text{ع}_{\text{س ص}} = \frac{\sum \text{س}_{\text{ر}}\text{ص}_{\text{ر}} - \dfrac{\sum \text{س}_{\text{ر}} \sum \text{ص}_{\text{ر}}}{\text{ن}}}{\text{ن} - ١}$$

$$= ٣٨,٥$$

$$\text{ع}^{٢}_{\text{س}} = \frac{\sum \text{س}_{\text{ر}}^{٢} - \dfrac{(\sum \text{س}_{\text{ر}})^{٢}}{\text{ن}}}{\text{ن} - ١}$$

$$= ١٤$$

$$\therefore \text{ع}_{\text{س}} = ٣,٧٤٢$$

$$\bar{\text{س}} = ١٥$$

$$\bar{\text{ص}} = ٣٦$$

$$\text{ب} = \frac{\text{ع}_{\text{س ص}}}{\text{ع}^{٢}_{\text{س}}} \qquad (١١)$$

= ٣٨,٥ / ١٤

ب = ٢,٧٥

أ = صَ - ب سَ (٩)

= ٣٦ - ٢,٧٥ × ١٥

أ = - ٥,٢٥

وبالتعويض فى المعادلة (٣) تكون :

صَ ر = -٥,٢٥ + ٢,٧٥س ر

تقدير الربح إذا بلغت تكلفة الدعاية ١٨ ألف ريال هو :

صَ ر = - ٥,٢٥ + ٢,٧٥ × ١٨

= ٤٤,٢٥ ألف ريال

= ٤٤٢٥٠ ريالاً

البرنامج التالى يقوم بإيجاد معادلة الانحدار الخطى البسيط. البيانات المستخدمة بالبرنامج والبيانات الواردة فى المثال (١١,١) السابق. باستخدام المعادلة :

$$Y = A + BX$$

وباعتبار أن

الوسط الحسابى للمتغير المستقل = X_3

الوسط الحسابى للمتغير التابع = Y_3

مجموع مربعات المتغير المستقل = X_2

مجموع المتغير المستقل = X_1

وبالتالى :-

$B = E/F$

التغاير = $E = (S - X_1 Y_1/N) / (N - 1)$

تباين المتغير المستقل = F

$A = Y_3 - BX_3$

```
10  REM برنامج لايجاد معادله الانحدار الخطى البسيط
20  DIM X(7),Y(7)
30   X1=0
40   Y1=0
50   X2=0
60   Y2=0
70  READ N,V
80  FOR I=1 TO N
90   READ X(I),Y(I)
100  X1=X1+X(I)
110  Y1=Y1+Y(I)
120  X2=X2+X(I)*X(I)
130  Y2=Y2+Y(I)*Y(I)
140  S=S+X(I)*Y(I)
150 NEXT I
160 E=(S-X1*Y1/N)/(N-1)
170 F=(X2-X1*X1/N)/(N-1)
180 F1=SQR(F)
190 X3=X1/N    REM MEAN OF X
200 Y3=Y1/N    REM MEAN OF Y
210 B=E/F
220 A=Y3-B*X3
230 R=A+B*V
240 REM  PRINTING SECTION
245 PRINT USING 330
250 PRINT USING 320
260 PRINT USING 330
270 FOR I=1 TO N
280  PRINT USING 340,X(I)*Y(I),Y(I)*Y(I),X(I)*X(I),Y(I),X(I),I
290 NEXT I
300 PRINT USING 330
310 PRINT USING 350,S,Y2,X2,Y1,X1
315 PRINT USING 330
320 :      س ر ص ر      ص ر٢      س ر٢      ص ر      س ر      رقم
330 :  ----------------------------------------------------------------
340 :       ####       ####      ####       ##       ##        ##
350 :      #####      #####     #####      ###      ###    المجموع
360 PRINT
370 PRINT , A;'= ا'
380 PRINT , B;'= ب'
390 PRINT
460 DATA 7,18,10,25,15,35,14,30,16,40,11,24,20,50,19,48
470 END
```

المخرجات

س ر ص ر	ص ر٢	س ر٢	ص ر	س ر	رقم
250	625	100	25	10	1
525	1225	225	35	15	2
420	900	196	30	14	3
640	1600	256	40	16	4
264	576	121	24	11	5
1000	2500	400	50	20	6
912	2304	361	48	19	7
4011	9730	1659	252	105	المجموع

```
-5.25   = ا
 2.75   = ب
```

۳ - خصائص معادلة الانحدار الخطى البسيط :

(أ) المعدل العام :

ص̂ ر = أ + ب س ر (۳)

أ = ص̄ - ب س̄ (۹)

بتعويض (۹) فى (۳) تكون :

ص̂ ر = ص̄ - ب س̄ + ب س ر

ص̂ ر = ص̄ + ب (س ر - س̄) (۱۲)

فإذا كانت

س ۱ = س ۲ = س ۳ = ... = س ر = ... س ن

فإن :

ص̂ ر = ص̄

أى أن ص̂ ر هى المعدل العام (الوسط الحسابى) إذا كانت س ر ليست متغيرة.

(ب) خط الانحدار يمر بالنقطة (س̄ ، ص̄) :

ص̂ ر = ص̄ + ب (س ر - س̄) (۱۲)

إذا كانت س ر = س̄ فإن :

ص̂ ر = ص̄

أى أن الخط يمر بنقطة الوسطين (س̄ ، ص̄). ففى المثال السابق مثلاً كانت المعادلة هى :

ص̂ ر = - ٥,٢٥ + ٢,٧٥ س ر

فإذا كانت س ر = س̄ = ١٥

فإن ص̂ ر = - ٥,٢٥ + ٢,٧٥ × ١٥

= ٣٦ وهى ص̄

(جـ) مربع الارتباط الخطى هو المقياس لدقة التقدير :

تباين المتغير التابع هو :

$$ع^2_{ص} = \frac{\sum (ص_ر - \bar{ص})^2}{ن - 1} \quad (13)$$

وهذا يعنى أن مجموع مربعات الانحرافات الكلى (م م ك) للمتغير التابع هو :

$$م\,م\,ك = \sum (ص_ر - \bar{ص})^2 \quad (14)$$

أما تباين التقدير ($\hat{ص}_ر$) فهو :

$$ع^2_{\hat{ص}_ر} = \frac{\sum (\hat{ص}_ر - \bar{ص})^2}{ن - 1} \quad (15)$$

إذاً فمجموع مربعات الانحرافات بسبب الانحدار (م م ر) هو :

$$م\,م\,ر = \sum (\hat{ص}_ر - \bar{ص})^2 .... \quad (16)$$

هذا ، ويلاحظ من المعادلة (٥)، والمعادلة (١٤)، والمعادلة (١٦) أن :

$$م\,م\,ك = م\,م\,ر + م\,م\,خ$$

أى أن :

$$\sum (ص_ر - \bar{ص})^2 = \sum (\hat{ص}_ر - \bar{ص})^2 + \sum (ص_ر - \hat{ص}_ر)^2 \quad (17)$$

وعليه فنسبة م م ر إلى م م ك هى نسبة التغيرات التى تفسرها (تكشفها) المعادلة. ويطلق على النسبة المئوية التى تفسرها معادلة الانحدار بمعامل التحديد (Coefficient of Determination) وهى قياس لمستوى دقة المعادلة ؛ لأنها تمثل عدد النقاط الواقعة على الخط من كل مائة نقطة. وبذلك يمكن تفسير معامل التحديد بأنه يساوى $\frac{م\,م\,ر}{م\,م\,ك}$

أى أن :

$$\text{معامل التحديد} = \frac{\sum (\hat{\text{ص}}_{\text{ر}} - \bar{\text{ص}})^{٢}}{\sum (\text{ص}_{\text{ر}} - \bar{\text{ص}})^{٢}} \times ١٠٠ \qquad (١٨)$$

إلا أن :

$$\hat{\text{ص}}_{\text{ر}} = \bar{\text{ص}} + \text{ب} (\text{س}_{\text{ر}} - \bar{\text{س}}) \qquad (١٢)$$

وبتعويض (١٢) فى (١٨) يكون :

$$\frac{\text{معامل التحديد}}{١٠٠} = \frac{(\bar{\text{ص}} + \text{ب} (\text{س}_{\text{ر}} - \bar{\text{س}}) - \bar{\text{ص}})^{٢}}{\sum (\text{ص}_{\text{ر}} - \bar{\text{ص}})^{٢}} \qquad (١٩)$$

$$= \frac{\sum (\text{ب} (\text{س}_{\text{ر}} - \bar{\text{س}}))^{٢}}{\sum (\text{ص}_{\text{ر}} - \bar{\text{ص}})^{٢}}$$

$$= \frac{\text{ب}^{٢} \sum (\text{س}_{\text{ر}} - \bar{\text{س}})^{٢}}{\sum (\text{ص}_{\text{ر}} - \bar{\text{ص}})^{٢}}$$

$$= \frac{\text{ب}^{٢} \text{ع}^{٢}_{\text{س}}}{\text{ع}^{٢}_{\text{ص}}} \qquad (٢٠)$$

وبالتعويض عن قيمة (ب) الواردة فى المعادلة (١١)

على النحو التالى :

$$\text{ب} = \frac{\text{ع}_{\text{س ص}}}{\text{ع}^{٢}_{\text{س}}} \qquad (١١)$$

يصبح معامل التحديد هو :

$$\left(\frac{\text{ع}_{\text{س ص}}}{\text{ع}^{٢}_{\text{س}}}\right)^{٢} \times \frac{\text{ع}^{٢}_{\text{س}}}{\text{ع}^{٢}_{\text{ص}}} \times ١٠٠$$

أى أن :

معامل التحديد

$$= \frac{\left(\text{ع}_{\text{س ص}}\right)^{٢}}{\text{ع}_{\text{س}}^{٢} \times \text{ع}_{\text{ص}}^{٢}} \times ١٠٠$$

$$= \frac{\text{ع}_{\text{س ص}}}{\text{ع}_{\text{س}}\,\text{ع}_{\text{ص}}} \times \frac{\text{ع}_{\text{س ص}}}{\text{ع}_{\text{س}}\,\text{ع}_{\text{ص}}} \times ١٠٠ \qquad (٢١)$$

هذا ويتضح من (٢١) أن :

معامل التحديد = ر × ر × ١٠٠ (٢٢)

= ١٠٠ $\text{ر}^{٢}$ (٢٣)

= ١٠٠ × مربع الارتباط الخطى .

مثال (٢ ,١١) :

أوجد معامل التحديد لخط الانحدار الوارد فى المثال (١) .

الحل :

$\text{ع}_{\text{س}}$ = ٣,٧٤٢

$\text{ع}_{\text{ص}}$ = ١٠,٤٧٢

$\text{ع}_{\text{س ص}}$ = ٣٨,٥

$$\text{ر} = \frac{\text{ع}_{\text{س ص}}}{\text{ع}_{\text{س}}\,\text{ع}_{\text{ص}}}$$

$$= \frac{٣٨,٥}{٣,٧٤٢ \times ١٠,٤٧٢}$$

= ٠,٩٨٢

معامل التحديد = ١٠٠ $\text{ر}^{٢}$ (٢٣)

= ١٠٠ × $(٠,٩٨٢)^{٢}$

= ٩٦,٤٪

بمعنى أن حوالى ٩٦ من كل مائة نقطة تكون على خط الانحدار.

البرنـامج التالى يقوم بحساب معامل التحديد لخط الانحدار الوارد فى المثال (١١،١) السابق ، اعتماداً على أن معامل التحديد هو مربع الارتباط الخطى حيث الارتباط الخطى هو :

$$R = \frac{E}{FG}$$

E = التغاير

F = الانحراف المعيارى للمتغير المستقل

G = الانحدار المعيارى للمتغير التابع

```
10  REM   برنامج لايجاد معامل التحديد
20  N=7
30  S=4011
40  X1=105          REM  SUM OF X
50  X2=1659         REM  SUM OF X SQUARE
60  Y1=252          REM  SUM OF Y
70  Y2=9730         REM  SUM OF Y SQUARE
80  E=(S-X1*Y1/N)/(N-1)
90  F=SQR((X2-X1*X1/N)/(N-1))
100 G=SQR((Y2-Y1*Y1/N)/(N-1))
110 R=E/(F*G)
120 PRINT ,E;'= عس ص'
130 PRINT
140 PRINT ,F;'= عس'
150 PRINT
160 PRINT ,G;'= عص'
170 PRINT
180 PRINT ,R*R;'= اذن معامل التحديد'
190 PRINT
200 END
```

المخرجات

```
38.5        = عس ص
3.741657    = عس
10.47218    = عص
.965426     = اذن معامل التحديد
```

٤ - انحرافات التقديرات:

(أ) الخطأ المعيارى لخط الانحدار (ع) :

تسمى الإحصائية $\frac{م م خ}{ن - ٢}$ بتباين خط الانحدار ($ع^٢$) أو تباين الخطأ (RESIDUAL).

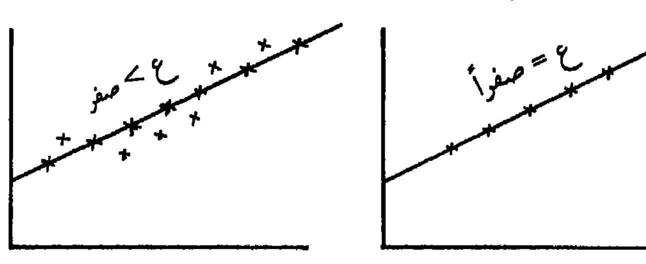
أى أن :

$$ع^٢ = \frac{\sum (ص_ر - \hat{ص}_ر)^٢}{ن - ٢} \qquad (٢٤)$$

بينما يسمى الجذر التربيعى الموجب لتباين الخطأ بالخطأ المعيارى للتقدير .
إذاً فالخطأ المعيارى لتقدير ص هو :

$$ع = \sqrt{\frac{\sum (ص_ر - \hat{ص}_ر)^٢}{ن - ٢}} \qquad (٢٥)$$

$$= \sqrt{\frac{\sum (ص_ر - أ \sum ص_ر - ب \sum س_ر ص_ر)^٢}{ن - ٢}}$$



شكل (٣): الخطأ المعيارى للتقدير ع ≥ صفر

هذا ، ولقد ورد فى المعادلة (١٧) أن :

$$م م ك = م م ر + م م خ \qquad (١٧)$$

إذاً :

$$م م خ = م م ك - م م ر \qquad (٢٦)$$

$$م م خ = \sum (ص_ر - \bar{ص})^٢ - \sum (\hat{ص}_ر - \bar{ص})^٢ \qquad (٢٧)$$

أما الخطأ المعيارى للميل (بَ) فهو :

$$ع_ب = \frac{ع}{\sqrt{\sum (س_ر - \bar{س})^2}} \qquad (٢٨)$$

بينما يعرف الخطأ المعيارى للمعلم أ بأنه :

$$ع_أ = \frac{ع}{\sqrt{\frac{ن \sum (س_ر - \bar{س})^2}{\sum س_ر^2}}} \qquad (٢٩)$$

وأما انحراف صَ$_ر$ المعروف باسم وسط مربع الخطأ (و م خ) فهو :

$$و م خ = ع \left[\frac{١}{ن} + \frac{(س_ر - \bar{س})^2}{\sum (س_ر - \bar{س})^2} \right]^{١/٢} \qquad (٣٠)$$

ففترة الثقة عند أى قيمة (س$_ر$) للقيمة التقديرية صَ$_ر$ بمستوى معنوية محدد هى :

$$صَ_ر \pm ت ع \sqrt{\frac{١}{ن} + \frac{(س_ر - \bar{س})^2}{\sum (س_ر - \bar{س})^2}} \qquad (٣١)$$

حيث ت هى القيمة المستخرجة من جدول توزيع ت (t) على (ن - ٢) درجات حرية .

٥ ـ الانحدار الثنائى :

يسمى الانحدار بالانحدار المتعدد إذا كان المتغير التابع (ص$_ر$) يعتمد على عدة متغيرات مستقلة (س$_{١ر}$ ، س$_{٢ر}$، س$_{٣ر}$ ، ٠٠٠٠٠) . أما المتغير الثنائى فهو أحد أنواع الانحدار المتعدد، فنموذجه يتكون من متغيرين مستقلين فقط، أى أنه على النحو الآتى :

$$صَ_ر = أ + ب س_{١ر} + جـ س_{٢ر} \qquad (٣٢)$$

ولإيجاد قيم المعالم أ ، ب ، جـ لابد من استخدام ثلاث معادلات طبيعية على النحو التالى :

$$\sum \text{ص}_{\text{ر}} = \text{أ ن} + \text{ب} \sum \text{س}_{\text{١ر}} + \text{جـ} \sum \text{س}_{\text{٢ر}} \qquad (٣٣)$$

$$\sum \text{س}_{\text{١ر}}\,\text{ص}_{\text{ر}} = \text{أ} \sum \text{س}_{\text{١ر}} + \text{ب} \sum \text{س}^{٢}_{\text{١ر}} + \text{جـ} \sum \text{س}_{\text{١ر}}\,\text{س}_{\text{٢ر}} \qquad (٣٤)$$

$$\sum \text{س}_{\text{٢ر}}\,\text{ص}_{\text{ر}} = \text{أ} \sum \text{س}_{\text{٢ر}} + \text{ب} \sum \text{س}_{\text{١ر}}\,\text{س}_{\text{٢ر}} + \text{جـ} \sum \text{س}^{٢}_{\text{٢ر}} \qquad (٣٥)$$

مثال (٣, ١١) :

استخدم البيانات التالية لإيجاد معادلة انحدار الدخل ($\text{ص}_{\text{ر}}$) على سنوات الخبرة ($\text{س}_{\text{١ر}}$) وعدد سنوات الدراسة ($\text{س}_{\text{٢ر}}$) والبيانات هى :

الرقم	الدخل الشهرى بآلاف الريالات	سنوات الخبرة	عدد سنوات الدراسة
١	٣	٢٣	٩
٢	٦	٢٨	١٢
٣	٥	٢١	١٢
٤	٨	٢٣	٢٢
٥	١٠	٣٠	١٦
٦	١٥	٣٢	١٨
٧	٩	٢٥	٢٣
المجموع	٥٦	١٨٢	١١٢

الحل

ص ر	س ١ر	س ٢ر	س ١ر ص ر	س ١ر ٢	س ١ر س ٢ر	س ٢ر ص ر	س ٢ر ٢
٣	٢٣	٩	٦٩	٥٢٩	٢٠٧	٢٧	٨١
٦	٢٨	١٢	١٦٨	٧٨٤	٣٣٦	٧٢	١١٤
٥	٢١	١٢	١٠٥	٤٤١	٢٥٢	٦٠	١٤٤
٨	٢٣	٢٢	١٨٤	٥٢٩	٥٠٦	١٧٦	٤٨٤
١٠	٣٠	١٦	٣٠٠	٩٠٠	٤٨٠	١٦٠	٢٥٦
١٥	٣٢	١٨	٤٨٠	١٠٢٤	٥٧٦	٢٧٠	٣٢٤
٩	٢٥	٢٣	٢٢٥	٦٢٥	٥٧٥	٢٠٧	٥٢٩
٥٦	١٨٢	١١٢	١٥٣١	٤٨٣٢	٢٩٣٢	٩٧٢	١٩٦٢

∑ ص ر = أن + ب∑ س ١ر + جـ∑ س ٢ر (٣٣)

∑ س ١ر ص ر = أ∑ س ١ر + ب∑ س ١ر ٢ + جـ∑ س ١ر س ٢ر (٣٤)

∑ س ٢ر ص ر = أ∑ س ٢ر + ب ∑ س ١ر س ٢ر + جـ∑ س ٢ر ٢ (٣٥)

بالتعويض فى المعادلات السابقة من البيانات بالجدول :

٥٦ = ٧ أ + ١٨٢ ب + ١١٢ جـ (٣٣)

١٥٣١ = ١٨٢ أ + ٤٨٣٢ ب + ٢٩٣٢ جـ (٣٤)

٩٧٢ = ١١٢ أ + ٢٩٣٢ ب + ١٩٦٢ جـ (٣٥)

بضرب المعادلة (٢٣) × ٢٦ :

١٤٥٦ = ١٨٢ أ + ٤٧٣٢ ب + ٢٩١٢ جـ (٣٣)

١٥٣١ = ١٨٢ أ + ٤٨٣٢ ب + ٢٩٣٢ جـ (٣٤)

∴ ٧٥ = ١٠٠ ب + ٢٠ جـ (٣٦)

بضرب (٣٣) × ١٦ :

٩٧٢ = ١١٢ أ + ٢٩٣٢ ب + ١٩٦٢ جـ (٣٥)

٨٩٦ = ١١٢ أ + ٢٩١٢ ب + ١٧٩٢ جـ (٣٣)

∴ ٧٦ = ٢٠ ب + ١٧٠ جـ (٣٧)

بحل المعادلتين (٣٦) و (٣٧) آنياً تكون :

٣٨٠ = ١٠٠ ب + ٨٥٠ جـ (٣٧)

٧٥ = ١٠٠ ب + ٢٠ جـ (٣٦)

∴ ٣٠٥ = ٨٣٠ جـ

جـ = ٠,٣٦٧

ب = ٠,٦٨٠

أ = - ١٥,٥٥

ومن ثم تصبح معادلة الانحدار الثنائى هى :

صَ ر = - ١٥,٥٥ + ٠,٦٨ س١ر + ٠,٣٦٧ س٢ر

ويبدو جلياً من الوارد سابقاً أن التعامل مع الانحدار المتعدد ليس سهلاً، خاصة إذا كان عدد المتغيرات كبيراً . ولهذا السبب تستخدم طريقة المصفوفات الواردة بعد، والتى تستخدم فيها الحاسبات الآلية كما سوف يرد فيها بعد.

٦ ـ الانحدار بالمصفوفات :

يسمى النموذج :

ص ر = أ + ب س ر (٢)

بالنموذج غير المركزى، بينما يسمى النموذج :

ص ر = أ + ب (س ر - سَ) (٣٨)

بالنمـوذج المـركزى. وتكون قيمة المعلم الأول أ = صَ فى حالة النموذج المركزى، بينما تتساوى قيمة ب فى النموذجين.

وسواء كان النموذج مركزياً أو غير مركزى، أو كان النموذج بسيطاً أو خطياً متعدداً ، فمن الممكن عرضه على النحو التالى :

$$\underline{\text{ص}} = \underline{\text{س}}\ \underline{\text{ب}} + \underline{\text{خ}} \qquad (٣٩)$$

وذلك باعتبار أن :

$\underline{\text{ص}}$ = متجهاً عمودياً من الرتبة ن × ١

$\underline{\text{س}}$ = مصفوفة المتغيرات المستقلة، وهى من الرتبة ث × ن حيث ث هى عدد المعالم المطلوب تقديرها. كما يجب الفصل هنا بين النموذج المركزى وغير المركزى إذا أن $\underline{\text{س}}$ تمثل المتغيرات المستقلة فى حالة النموذج غير المركزى، وتمثل انحرافات المتغيرات فى حالة النموذج المركزى.

$\underline{\text{ب}}$ = متجها عمودياً من الرتبة ث × ١ .

إذاً فالنموذج الخاص بالقيم التقديرية يكون على النحو التالى :

$$\underline{\hat{\text{ص}}} = \underline{\text{س}}\ \underline{\hat{\text{ب}}} \qquad (٤٠)$$

يمكن تمثيله بالمصفوفات على النحو الآتى :

$$\begin{pmatrix} \text{ص}_{١} \\ \text{ص}_{٢} \\ \text{ص}_{٣} \\ \vdots \\ \vdots \\ \text{ص}_{ن} \end{pmatrix} = \begin{pmatrix} ١ & \text{س}_{١١} & \text{س}_{١٢} & \cdots \\ ١ & \text{س}_{٢١} & \text{س}_{٢٢} & \\ ١ & \text{س}_{٣١} & \text{س}_{٣٢} & \\ \vdots & \vdots & \vdots & \\ \vdots & \vdots & \vdots & \\ ١ & \text{س}_{ن١} & \text{س}_{ن١} & \end{pmatrix} \begin{pmatrix} \text{أ} \\ \text{ب}_{١} \\ \text{ب}_{٢} \\ \vdots \\ \vdots \end{pmatrix} \qquad (٤١)$$

وبضرب طرفى المعادلة (٤٠) فى $\underline{\text{س}}'$ (مدور المصفوفة $\underline{\text{س}}$) تكون :

$$\underline{\text{س}}'\ \underline{\text{ص}} = \underline{\text{س}}'\ \underline{\text{س}}\ \underline{\text{ب}} \qquad (٤٢)$$

وبالضرب المسبق لطرفى المعادلة السابقة فى مقلوب $\underline{\text{س}}'\ \underline{\text{س}}$ يمكن الحصول على النموذج الخطى لأى عدد من المتغيرات المستقلة. وذلك على النحو التالى :

$$(\underline{\text{س}}'\ \underline{\text{س}})^{-١}\ \underline{\text{س}}'\ \underline{\text{ص}} = (\underline{\text{س}}'\ \underline{\text{س}})^{-١}\ (\underline{\text{س}}'\ \underline{\text{س}})\ \underline{\text{ب}}$$

$$\therefore\ \underline{\text{ب}} = (\underline{\text{س}}'\ \underline{\text{س}})^{-١}\ \underline{\text{س}}'\ \underline{\text{ص}}$$

بيد أن إيجاد النظير الضربي يكون مطولاً إذا كان عدد المتغيرات المستقلة كبيراً . لذلك تستخدم الحاسبات الآلية كثيراً في هذا المجال.

أ ـ الانحدار الخطي البسيط بالمصفوفات :

$$\underline{ص} = \underline{س}\ \underline{ب} \qquad (٤٠)$$

$$\underline{س'} = \begin{pmatrix} ١ & ١ & ١ & \cdots\cdots & ١ \\ س_١ & س_٢ & س_٣ & & س_ن \end{pmatrix}$$

$$\underline{ب'} = (أ \quad ب)$$

$$\underline{ص'} = (ص_١ \; ص_٢ \; ص_٣ \; \cdots\cdots ، ص_ن)$$

ومن ثم فإن :

$$\begin{pmatrix} ص_١ \\ ص_٢ \\ ص_٣ \\ \vdots \\ \vdots \\ ص_ن \end{pmatrix} = \begin{pmatrix} ١ & س_١ \\ ١ & س_٢ \\ ١ & س_٣ \\ \vdots & \vdots \\ \vdots & \vdots \\ ١ & س_ن \end{pmatrix} \begin{pmatrix} أ \\ \\ ب \end{pmatrix}$$

ويتضح من المعادلة السابقة أن :

$$ص_١ = أ + ب\ س_١$$

$$ص_٢ = أ + ب\ س_٢$$

$$ص_٣ = أ + ب\ س_٣$$

$$\vdots$$

$$ص_ن = أ + ب\ س_ن$$

$$\underline{س'}\ \underline{س} = \begin{pmatrix} ١ & ١ & ١ & \cdots & ١ \\ س_١ & س_٢ & س_٣ & \cdots & س_ن \end{pmatrix} \begin{pmatrix} ١ & س_١ \\ ١ & س_٢ \\ ١ & س_٣ \\ \vdots & \vdots \\ ١ & س_ن \end{pmatrix}$$

$$= \begin{bmatrix} ن & \sum س_ر \\ \sum س_ر & \sum س_ر^٢ \end{bmatrix} \qquad (٤٤)$$

إذاً سَ س مصفوفة متماثلة على قطرها الرئيسى عدد المتغيرات ومجموع مربعاتها، وعلى القطر الثانوى مجموع المتغيرات.

$$\underline{س'}\ \underline{ص} = \begin{pmatrix} ١ & ١ & ١ & \cdots & ١ \\ س_١ & س_٢ & س_٣ & & س_ن \end{pmatrix} \begin{pmatrix} ص_١ \\ ص_٢ \\ ص_٣ \\ \\ ص_ن \end{pmatrix}$$

$$\therefore \underline{س'}\ \underline{ص} = \begin{pmatrix} \sum ص_ر \\ \sum س_ر ص_ر \end{pmatrix} \qquad (٤٥)$$

فهى إذاً متجه عمودى من الرتبة الثانية، ويتكون من مجموع المتغيرات التابعة ومجموع مضاريب تلك المتغيرات فى المتغيرات المستقلة.

$$\underline{ب} = (\underline{س'}\ \underline{س})^{-١}\ \underline{س'}\ \underline{ص} \qquad (٤٣)$$

$$(\underline{س'}\ \underline{س})^{-١} = \begin{bmatrix} ن & \sum س_ر \\ \sum س_ر & \sum س_ر^٢ \end{bmatrix}^{-١}$$

$$= \frac{1}{ن\left(\sum س_ر^{2} - \frac{(\sum س_ر)^{2}}{ن}\right)} \begin{bmatrix} \sum س_ر^{2} & -\sum س_ر \\ -\sum س_ر & ن \end{bmatrix}'$$

$$= \frac{1}{ن\left(\sum س_ر^{2} - \frac{(\sum س_ر)^{2}}{ن}\right)} \begin{bmatrix} \sum س_ر^{2} & -\sum س_ر \\ -\sum س_ر & ن \end{bmatrix} \qquad (٤٦)$$

$$\therefore (\underline{س}'\,\underline{س})^{-1}\,\underline{س}'\,\underline{ص} = \frac{1}{ن\left(\sum س_ر^{2} - \frac{(\sum س_ر)^{2}}{ن}\right)} \begin{bmatrix} \sum س_ر^{2} & -\sum س_ر \\ -\sum س_ر & ن \end{bmatrix} \begin{bmatrix} \sum ص_ر \\ \sum س_ر ص_ر \end{bmatrix}$$

$$= \frac{1}{ن\left(\sum س_ر^{2} - \frac{(\sum س_ر)^{2}}{ن}\right)} \begin{bmatrix} \sum س_ر^{2} \sum ص_ر - \sum س_ر \sum س_ر ص_ر \\ -\sum س_ر \sum ص_ر + ن \sum س_ر ص_ر \end{bmatrix} \qquad (٤٧)$$

فالمقدار الأول يمكن عرضه على النحو التالى :

$$\frac{\sum س_ر^{2} \sum ص_ر - \sum س_ر \sum س_ر ص_ر}{ن\left[\sum س_ر^{2} - \frac{(\sum س_ر)^{2}}{ن}\right]}$$

$$= \frac{\sum س_ر^{2} \sum ص_ر - \sum ص_ر \frac{(\sum س_ر)^{2}}{ن} + \sum ص_ر \frac{(\sum س_ر)^{2}}{ن} - \sum س_ر \sum س_ر ص_ر}{ن\left[\sum س_ر^{2} - \frac{(\sum س_ر)^{2}}{ن}\right]}$$

$$= \frac{\sum ص_ر \left[\sum س_ر^{٢} - \frac{(\sum س_ر)^{٢}}{ن}\right] - \sum س_ر \left[\sum س_ر ص_ر - \frac{\sum س_ر \sum ص_ر}{ن}\right]}{ن \left[\sum س_ر^{٢} - \frac{(\sum س_ر)^{٢}}{ن}\right]}$$

$$= \bar{ص} - \bar{س} \; \frac{\left[\sum س_ر ص_ر - \frac{\sum س_ر \sum ص_ر}{ن}\right]}{\sum س_ر^{٢} - \frac{(\sum س_ر)^{٢}}{ن}}$$

$$= \bar{ص} - ب\,\bar{س}$$

وأما المقدار الثانى فهو :

$$ب = \frac{\sum س_ر ص_ر - \frac{\sum س_ر \; \sum ص_ر}{ن}}{\left(\sum س_ر^{٢} - \frac{(\sum س_ر)^{٢}}{ن}\right)}$$

وبالتالى فإن :

$$(\underline{س}'\,\underline{س})^{-1}\,\underline{س}'\,\underline{ص} = \underline{ب} \qquad (٤٣)$$

حيث $\underline{ب}$ هى متجه المعالم، وأولها أ وثانيها ب. هذا ويمكن توضيح ذلك بتطبيق طريقة المصفوفات على المثال رقم (١) ، ومن ثم مقارنة النتائج فى المثالين.

مثال (٤ , ١١):

استخدم البيانات الواردة فى مثال (١) لإيجاد معادلة الانحدار الخطى البسيط للربح على الدعاية . والبيانات هى :

الرقم	الدعاية ($س_ر$) بآلاف الريالات	الربح ($ص_ر$) بآلاف الريالات
١	١٠	٢٥
٢	١٥	٣٥
٣	١٤	٣٠
٤	١٦	٤٠
٥	١١	٢٤
٦	٢٠	٥٠
٧	١٩	٤٨
المجموع	١٠٥	٢٥٢

الحل :

$$\underline{س}'\underline{س} = \begin{pmatrix} ١ & ١ & ١ & ١ & ١ & ١ & ١ \\ ١٩ & ٢٠ & ١١ & ١٦ & ١٤ & ١٥ & ١٠ \end{pmatrix} \begin{pmatrix} ١٠ & ١ \\ ١٥ & ١ \\ ١٤ & ١ \\ ١٦ & ١ \\ ١١ & ١ \\ ٢٠ & ١ \\ ١٩ & ١ \end{pmatrix}$$

$$= \begin{pmatrix} ١٠٥ & ٧ \\ ١٦٥٩ & ١٠٥ \end{pmatrix}$$

ويلاحظ هنا بالمقارنة مع حل المثال رقم (١) أن :

$ن = ٧$

$\sum س_ر = ١٠٥$

$\sum س_ر^٢ = ١٦٥٩$

$$\underline{س}'\,\underline{ص} = \begin{pmatrix} ١ & ١ & ١ & ١ & ١ & ١ & ١ \\ ١٠ & ١٥ & ١٤ & ١٦ & ١١ & ٢٠ & ١٩٠ \end{pmatrix} \begin{pmatrix} ٢٥ \\ ٣٥ \\ ٣٠ \\ ٤٠ \\ ٢٤ \\ ٥٠ \\ ٤٨ \end{pmatrix}$$

$$= \begin{pmatrix} ٢٥٢ \\ ٤٠١١ \end{pmatrix}$$

ومن المثال رقم (١) اتضح أن

$$\sum ص_ر = ٢٥٢$$

$$\sum س_ر ص_ر = ٤٠١١$$

$$(\underline{س}'\,\underline{س})^{-١} = \begin{bmatrix} ٧ & ١٠٥ \\ ١٠٥ & ١٦٥٩ \end{bmatrix}^{-١}$$

$$= \frac{١}{٥٨٨} \begin{bmatrix} ١٦٥٩ & -١٠٥ \\ -١٠٥ & ٧ \end{bmatrix}$$

$$\underline{ب} = (\underline{س}'\,\underline{س})^{-١}\, \underline{س}'\,\underline{ص} \qquad (٤٣)$$

$$= \frac{١}{٥٨٨} \begin{bmatrix} ١٦٥٩ & -١٠٥ \\ -١٠٥ & ٧ \end{bmatrix} \begin{bmatrix} ٢٥٢ \\ ٤٠١١ \end{bmatrix}$$

$$= \begin{pmatrix} -٥,٢٥ \\ ٢,٧٥ \end{pmatrix}$$

$$\therefore أ = -٥,٢٥$$

$$ب = ٢,٧٥$$

ومعادلة الانحدار الخطى هى :

$$\hat{ص}_ر = -٥,٢٥ + ٢,٧٥\, س_ر$$

وهى نفس معادلة المثال رقم (١).

ب ـ استخدامات أخرى للمصفوفات السابقة :

يمكن إعادة صياغة المعادلة (٢٨) الخاصة بتباين المتغير ب على النحو التالى :

$$ع_{ب}^{٢} = \frac{ع^{٢}}{\sum (س_{ر} - \bar{س})^{٢}}$$

$$= \frac{ع^{٢}}{\sum س_{ر}^{٢} - \frac{(\sum س_{ر})^{٢}}{ن}} \qquad (٢٨)$$

والمعادلة (٢٩) الخاصة بتباين أ على النحو التالى :

$$ع_{أ}^{٢} = \frac{ع^{٢} \sum س_{ر}^{٢}}{ن \sum (س_{ر} - \bar{س})^{٢}}$$

$$= \frac{ع^{٢} \sum س_{ر}^{٢}}{ن \left(\sum س_{ر}^{٢} - \frac{(\sum س_{ر})^{٢}}{ن} \right)} \qquad (٢٩)$$

وأما تغاير المعلمين أ، ب فهو :

$$ع_{أب} = ع \; (أ، ب)$$

$$= ع \; (\bar{ص} - ب \bar{س} \text{ ، } ب)$$

بيد أن $ع \; (\bar{ص} \text{ ، } ب) =$ صفراً

$$\therefore ع_{أب} = ع \; (\bar{ص} ب - ب^{٢} \bar{س})$$

$$= ع \; (- ب^{٢} \bar{س})$$

$$= - \bar{س} \; ع \; (ب^{٢})$$

$$= \frac{- \bar{س} ع^{٢}}{\sum س_{ر}^{٢} - \frac{(\sum س_{ر})^{٢}}{ن}} \qquad (٤٨)$$

وبالتالى يمكن عرض مصفوفة تشتت المتغيرين على النحو الآتى :

$$\text{ع}_{\text{اب}} = \begin{bmatrix} \text{ع}^2_{\text{أ}} & \text{ع}_{\text{اب}} \\ \text{ع}_{\text{با}} & \text{ع}^2_{\text{ب}} \end{bmatrix}$$

$$= \text{ع}^2 \begin{bmatrix} \dfrac{\sum \text{س}^2_{\text{ر}}}{\text{ن}\left(\sum \text{س}^2_{\text{ر}} - \dfrac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}}\right)} & \dfrac{-\bar{\text{س}}}{\sum \text{س}^2_{\text{ر}} - \dfrac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}}} \\ \dfrac{-\bar{\text{س}}}{\sum \text{س}^2_{\text{ر}} - \dfrac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}}} & \dfrac{1}{\sum \text{س}^2_{\text{ر}} - \dfrac{(\sum \text{س}_{\text{ر}})^2}{\text{ن}}} \end{bmatrix} \qquad (٤٩)$$

ويلاحظ أن المعادلة المبينة بالمصفوفة (٤٦) هى نفس المصفوفة السابقة (٤٩) بعد ضربها فى ع^2 . إذاً فالمصفوفة المتماثلة $(\underline{\text{س}}' \, \underline{\text{س}})^{-1}$ تستخدم أيضاً كمصفوفة تشتت للمعالم بعد ضربها فى ع^2 التى سيتم تقديرها على ما يلى :

اتضح مما مضى أن :

$$\text{م م ر} = \sum (\hat{\text{ص}}_{\text{ر}} - \bar{\text{ص}})^2$$
$$= \text{ب}^2 \sum (\text{س}_{\text{ر}} - \bar{\text{س}})^2 \qquad (٢٠)$$

وهى خاصة بمجموع مربعات الانحدار بسبب ب وحدها دون أ . أما مجموع المربعات بسبب أ وحدها فيسمى معامل تصحيح الوسط (correction for Mean) وهو عبارة عن :

$$\text{م م أ} = \text{ن} \, \bar{\text{ص}}^2$$
$$= \frac{(\sum \text{ص}_{\text{ر}})^2}{\text{ن}} \qquad (٥٠)$$

وبالتالى فإن :

$$\text{م م ر} + \text{م م أ} = \text{ب}^{2}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2} + \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}} \qquad (51)$$

بيد أن :

$$\text{ب}^{2} = \frac{\left[\Sigma\,(\text{س}_{\text{ر}} - \bar{\text{س}})\,(\text{ص}_{\text{ر}} - \bar{\text{ص}})\right]^{2}}{\left[\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}\right]^{2}} \qquad (10)$$

$$\therefore \text{م م ر} + \text{م م أ} = \frac{\left[\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})\,(\text{ص}_{\text{ر}} - \bar{\text{ص}})\right]^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}}$$

$$= \frac{\left[\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}} - \frac{\Sigma \text{س}_{\text{ر}}\,\Sigma \text{ص}_{\text{ر}}}{\text{ن}}\right]^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}}$$

$$= \frac{(\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}})^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} - \frac{2\,\Sigma \text{س}_{\text{ر}}\,\Sigma \text{ص}_{\text{ر}}\,\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{(\Sigma \text{س}_{\text{ر}})^{2}(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}^{2}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}}$$

$$= \frac{(\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}})^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} - \frac{2\,\Sigma \text{س}_{\text{ر}}\,\Sigma \text{ص}_{\text{ر}}\,\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}}\left[\frac{(\Sigma \text{س}_{\text{ر}})^{2}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + 1\right]$$

$$= \frac{(\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}})^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} - \frac{2\,\Sigma \text{س}_{\text{ر}}\,\Sigma \text{ص}_{\text{ر}}\,\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}}\left[\frac{\frac{(\Sigma \text{س}_{\text{ر}})^{2}}{\text{ن}} + \Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}}\right]$$

$$= \frac{(\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}})^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} - \frac{2\,\Sigma \text{س}_{\text{ر}}\,\Sigma \text{ص}_{\text{ر}}\,\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}}\left[\frac{(\Sigma \text{س}_{\text{ر}})^{2}}{\text{ن}} + \text{س}_{\text{ر}}^{2} - \frac{(\Sigma \text{س}_{\text{ر}})^{2}}{\text{ن}}\right]$$

$$= \frac{(\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}})^{2}}{\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} - \frac{2\,\Sigma \text{س}_{\text{ر}}\,\Sigma \text{ص}_{\text{ر}}\,\Sigma \text{س}_{\text{ر}}\text{ص}_{\text{ر}}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} + \frac{\Sigma \text{س}_{\text{ر}}^{2}\,(\Sigma \text{ص}_{\text{ر}})^{2}}{\text{ن}\,\Sigma(\text{س}_{\text{ر}} - \bar{\text{س}})^{2}} \qquad (52)$$

ويلاحظ من (٤٥) والمعادلة (٤٧) أن :

$$\underline{ب'}\,\underline{س'}\,\underline{ص} =$$

$$\frac{1}{ن\sum(س_ر - \bar{س})^2}\begin{bmatrix} \sum س_ر^2 \sum ص_ر - \sum س_ر \sum س_ر ص_ر & -\sum س_ر \sum ص_ر + ن \sum س_ر ص_ر \end{bmatrix}\begin{bmatrix}\sum ص_ر \\ \sum س_ر ص_ر\end{bmatrix}$$

$$= \frac{1}{ن\sum(س_ر - \bar{س})^2}\Big[\sum س_ر^2 (\sum ص_ر)^2 - \sum س_ر \sum ص_ر \sum س_ر ص_ر - \sum س_ر \sum ص_ر \sum س_ر ص_ر + ن(\sum س_ر ص_ر)^2\Big]$$

$$\therefore \underline{ب'}\,\underline{س'}\,\underline{ص} = \frac{\sum س_ر^2 (\sum ص_ر)^2}{ن\sum(س_ر - \bar{س})^2} - \frac{2\sum س_ر \sum ص_ر \sum س_ر ص_ر}{ن\sum(س_ر - \bar{س})^2} + \frac{(\sum س_ر ص_ر)^2}{\sum(س_ر - \bar{س})^2} \qquad (٥٣)$$

وبمقارنة نتيجة المعادلة (٥٣) السابقة بنتيجة المعادلة (٥٢) يتضح أن :

$$\underline{ب'}\,\underline{س'}\,\underline{ص} = م\,م\,ر + م\,م\,أ \qquad (٥٤)$$

ومن ثم تصبح :

$$ر^2 = \frac{\underline{ب'}\,\underline{س'}\,\underline{ص} - م\,م\,أ}{\underline{ص'}\,\underline{ص} - م\,م\,أ} \qquad (٥٥)$$

إذاً :

$$ر^2 = \frac{\underline{ب'}\,\underline{س'}\,\underline{ص} - ن\,\bar{ص}^2}{\underline{ص'}\,\underline{ص} - ن\,\bar{ص}^2} \qquad (٥٦)$$

وبذلك يمكن الاستفادة من نتائج تلك المصفوفات لاختبار معنوية المعلم (ب) باستخدام ما يسمى جدول تحليل التباين (ANOVA) ، وذلك على النحو الآتى :

جدول تحليل التباين لاختبار الفرضية :

ف$_{٠}$: ب = صفراً
ف$_{١}$: ب ≠ صفراً

ولتقدير التباين ع^{٢}

مصدر التباين	مجموع المربعات	درجات الحرية (د . ج)	متوسط المربعات	إحصائية الاختبار ف (د، ن − د)
أ	ن صَ^{٢}	١		
الانحدار	م م ر = بَ سَ صَ − ن صَ^{٢}	ث − ١ = د	$\frac{\text{م م ر}}{\text{د}}$ = م ر	$\frac{\text{م ر}}{\text{ع}^{٢}}$
الخطأ (بالطرح)	م م خ = م م ك − م م ر	ن − ث + ١	$\frac{\text{م م خ}}{\text{ن − د − ١}}$ = ع^{٢}	
المجموع الكلى	صَ صَ − ن صَ^{٢} = م م ك	ن − ١		

مثال (٥ , ١١) :

استخدم حل المثال (٤) لإيجاد ما يلى :

١ ـ اختبار الفرضية ف : ب = صفراً
٢ ـ مصفوفة التشتت للمعلمين.
٣ ـ معامل التحديد.

الحل :

جدول تحليل التباين لاختبار الفرضية

ف$_{٠}$: ب = صفراً
ف$_{١}$: ب ≠ صفراً

ومن المثال السابق :

$$\text{ب} = \begin{pmatrix} -٥,٢٥ \\ ٢,٧٥ \end{pmatrix}$$

$$(\underline{\text{س}'}\,\underline{\text{س}})^{-١} = \frac{١}{٥٨٨}\begin{bmatrix} ١٦٥٩ & -١٠٥ \\ -١٠٥ & ٧ \end{bmatrix}$$

$$\underline{\text{س}'}\,\underline{\text{ص}} = \begin{pmatrix} ٢٥٢ \\ ٤٠١١ \end{pmatrix}$$

$$\underline{\text{ص}'}\,\underline{\text{ص}} = \text{مج ص}_{\text{ر}}^{٢}$$

$$= ٩٧٣٠$$

$$\text{ن} = ٧$$

$$\bar{\text{ص}} = ٣٦$$

$$\text{م م ر} = \underline{\text{ب}'}\,\underline{\text{س}'}\,\underline{\text{ص}} - \text{ن}\,\bar{\text{ص}}^{٢}$$

$$= (-٥,٢٥ \quad ٢,٧٥)\begin{pmatrix} ٢٥٢ \\ ٤٠١١ \end{pmatrix} - ٧ \times ٣٦ \times ٣٦$$

$$= ٩٧٠٧,٢٥ - ٩٠٧٢$$

$$= ٦٣٥,٢٥٠$$

$$\text{م م ك} = \underline{\text{ص}'}\,\underline{\text{ص}} - \text{ن}\,\bar{\text{ص}}^{٢}$$

$$= ٩٧٣٠ - ٩٠٧٢$$

$$= ٦٥٨$$

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف ٥,١
الانحدار (ب)	٦٣٥,٢٥	١	٦٣٥,٢٥	١٣٩,٦١٥
الخطأ (بالطرح)	٢٢,٧٥	٥	٤,٥٥	
المجموع الكلى	٦٥٨	٦		

١ - اختبار الفرضية $\text{ف}_{٠}$: ب = صفراً

يتضح من جدول تحليل التباين أن إحصائية الاختبار تساوى ١٣٩,٦١٥ وهى جوهرية بمستوى معنوية ٩٥٪ مقارنة بتوزيع ف (١، ٥).

٢ - يمكن تقدير $\text{ع}^{٢}$ بمتوسط مربعات الخطأ، وبضرب $\text{ع}^{٢}$ فى مقلوب <u>سَ</u> <u>س</u> يمكن الحصول على مصفوفة التشتت التى يكون على قطرها تباينات المعالم، وحول القطر التغايرات. إذاً فهى :

$$\frac{\text{٤,٥٥}}{\text{٥٨٨}}\begin{bmatrix} \text{١٦٥٩} & \text{١٠٥ -} \\ \text{١٠٥ -} & \text{٧} \end{bmatrix}$$

٣ - معامل التحديد ($\text{ر}^{٢}$) :

$$\text{ر}^{٢} = \frac{\text{م م ر}}{\text{م م ك}} \qquad (\text{٥٥})$$

$$= \frac{\text{٦٣٥,٢٥}}{\text{٦٥٨}}$$

$$= \text{٠,٩٦٥}$$

وهى نفس نتيجة المثال (٢) تقريباً.

يتضح مما مضى أن طريقة المصفوفات فى الانحدار تتميز بما يلى :

١ - إمكانية تطبيق نفس المعادلة [<u>ب</u> = (<u>سَ</u> <u>س</u>)$^{-١}$ <u>سَ</u> ص] لتقدير المعالم لأى نموذج انحدار خطى مهما تكن المتغيرات.

٢ - سهولة تطبيق معادلاتها، خاصة لإيجاد معامل التحديد وتباينات وتغايرات المتغيرات؛ وذلك لأن بعض المصفوفات تستخدم لتنفيذ أكثر من مهمة واحدة.

٣ - سهولة تطبيقاتها بالحاسب الآلى، وبلغة البيسك؛ للاستفادة من سرعة ودقة الحاسبات خاصة وهناك دالة خاصة للمصفوفات فى لغة البيسك (MAT Function).

الانحدار المتعدد بالمصفوفات :

ورد فى أول وثانى خواص استخدام المصفوفات إمكانية استخدام نفس الأسلوب لأى عدد من المتغيرات. والمثال التالى هو إعادة للمثال (٣) الخاص بمتغيرين.

مثال (٦ر١١) :

استخـدم البيانات الواردة فى المثال (٣) لإيجاد معادلة الانحدار الخطى بالمصفوفات لتقدير الدخل من المدة (س١) وعدد سنوات الدراسة . والبيانات هى :

ص	س ١	س ٢
٣	٢٣	٩
٦	٢٨	١٢
٥	٢١	١٢
٨	٢٣	٢٢
١٠	٣٠	١٦
١٥	٣٢	١٨
٩	٢٥	٢٣
٥٦	١٨٢	١١٢

ثم أوجد ما يلى :

١ – اختبر الفرضية ف. : ب ١ = ب ٢= صفراً .

٢ – مصفوفة التشتت للمعالم .

٣ – معامل التحديد .

٤ – اختبر معنوية كل معلم بمستوى ٩٥ر.

الحل :

$$\underline{\text{ب}} = (\underline{\text{س}}'\,\underline{\text{س}})^{-1}\,\underline{\text{س}}'\,\underline{\text{ص}} \qquad (٤٣)$$

$$\underline{\text{ص}} = \underline{\text{س}}\,\underline{\text{ب}} \qquad (٤٠)$$

بتطبيق المعادلة (٤٠) تصبح المصفوفات على النحو الآتى :

$$\underline{\text{س ب}} = \underline{\text{ص}}$$

$$\begin{pmatrix} ٩ & ٢٣ & ١ \\ ١٢ & ٢٨ & ١ \\ ١٢ & ٢١ & ١ \\ ٢٢ & ٢٣ & ١ \\ ١٦ & ٣٠ & ١ \\ ١٨ & ٣٢ & ١ \\ ٢٣ & ٢٥ & ١ \end{pmatrix} = \begin{pmatrix} ٣ \\ ٦ \\ ٥ \\ ٨ \\ ١٠ \\ ١٥ \\ ٩ \end{pmatrix}$$

$$\underline{\text{س}}'\,\underline{\text{س}} = \begin{pmatrix} ١ & ١ & ١ & ١ & ١ & ١ & ١ \\ ٢٣ & ٢٨ & ٢١ & ٢٣ & ٣٠ & ٣٢ & ٢٥ \\ ٩ & ١٢ & ١٢ & ٢٢ & ١٦ & ١٨ & ٢٣ \end{pmatrix} \begin{pmatrix} ١ & ٢٣ & ٩ \\ ١ & ٢٨ & ١٢ \\ ١ & ٢١ & ١٢ \\ ١ & ٢٣ & ٢٢ \\ ١ & ٣٠ & ١٦ \\ ١ & ٣٢ & ١٨ \\ ١ & ٢٥ & ٢٣ \end{pmatrix}$$

$$= \begin{bmatrix} ٧ & ١٨٢ & ١١٢ \\ ١٨٢ & ٤٨٣٢ & ٢٩٣٢ \\ ١١٢ & ٢٩٣٢ & ١٩٦٢ \end{bmatrix}$$

$$\underline{\text{س}}'\,\underline{\text{ص}} = \begin{pmatrix} ١ & ١ & ١ & ١ & ١ & ١ & ١ \\ ٢٣ & ٢٨ & ٢١ & ٢٣ & ٣٠ & ٣٢ & ٢٥ \\ ٩ & ١٢ & ١٢ & ٢٢ & ١٦ & ١٨ & ٢٣ \end{pmatrix} \begin{pmatrix} ٣ \\ ٦ \\ ٥ \\ ٨ \\ ١٠ \\ ١٥ \\ ٩ \end{pmatrix}$$

$$\therefore \underline{\text{س}}'\,\underline{\text{ص}} = \begin{bmatrix} ٥٦ \\ ١٥٣١ \\ ٩٧٢ \end{bmatrix}$$

$$(\underline{\text{س}}'\,\underline{\text{س}})^{-١} = \begin{bmatrix} ٧ & ١٨٢ & ١١٢ \\ ١٨٢ & ٤٨٣٢ & ٢٩٣٢ \\ ١١٢ & ٢٩٣٢ & ١٩٦٢ \end{bmatrix}^{-١}$$

$$= \frac{١}{١١٦٢٠٠} \begin{bmatrix} ٨٨٣٧٦٠ & -٢٨٧٠٠ & -٧٥٦٠ \\ -٢٨٧٠٠ & ١١٩٠ & -١٤٠ \\ -٧٥٦٠ & -١٤٠ & ٧٠٠ \end{bmatrix}$$

$$\underline{ب} = (\underline{س}' \cdot \underline{س})^{-1} \; \underline{س}' \, \underline{ص} \qquad (٤٣)$$

$$\therefore \underline{ب} = \frac{١}{١١٦٢٠٠} \begin{bmatrix} ٨٨٣٧٦٠ & ٢٨٧٠٠ & -٧٥٦٠ \\ -٢٨٧٠٠ & ١١٩٠ & -١٤٠ \\ -٧٥٦٠ & -١٤٠ & ٧٠٠ \end{bmatrix} \begin{bmatrix} ٥٦ \\ ١٥٣١ \\ ٩٧٢ \end{bmatrix}$$

$$= \begin{bmatrix} -١٥٫٤٦٩ \\ ٠٫٦٧٦٥ \\ ٠٫٣٦٧٥ \end{bmatrix}$$

إذاً معادلة الانحدار هى :

$$\hat{ص}_{ر} = -١٥,٤٦٩ + ٠,٦٧٦٥ \, س_{١ر} + ٠,٣٦٧٥ \, س_{٢ر}$$

أما بقية التحليلات فتتم بناء على الخطوات التالية :

$$ن \, \bar{ص}^{٢} = \frac{ن(\sum ص_{ر})^{٢}}{ن^{٢}}$$

$$= \frac{٧ \times ٥٦ \times ٥٦}{٤٩}$$

$$= ٤٤٨$$

$$\underline{ب}' \, \underline{س}' \, \underline{ص} = \begin{pmatrix} -١٥,٤٦٩ & ٠,٦٧٦٥ & ٠,٣٦٧٥ \end{pmatrix} \begin{pmatrix} ٥٦ \\ ١٥٣١ \\ ٩٧٢ \end{pmatrix}$$

$$= ٥٢٦,٦٦٧٥$$

$$\underline{ص}' \, \underline{ص} = \begin{pmatrix} ٣ & ٦ & ٥ & ٨ & ١٠ & ١٥ & ٩ \end{pmatrix} \begin{pmatrix} ٣ \\ ٦ \\ ٥ \\ ٨ \\ ١٠ \\ ١٥ \\ ٩ \end{pmatrix}$$

= ٥٤٠

= $ص_ر^٢$

$\underline{ص} \underline{ص} - ن \bar{ص}^٢$ = ٥٤٠ ـ ٤٤٨

= ٩٢

$\underline{ب} \underline{س} \underline{ص} - ن \bar{ص}^٢$ = ٥٢٦,٦٦٧٥ - ٤٤٨

= ٧٨,٦٦٥

(١) وفيما يلي جدول تحليل التباين لاختبار الفرضية :

ف. : ب $_١$= ب $_٢$= صفراً

ف $_١$: هناك على الأقل ب $_١$ أو ب $_٢$ ≠ صفراً

المصدر	مجموع المربعات	درجات الحرية	متوسط المربعات	ف $_{٤،٢}$
الانحدار	٧٨,٦٦٧٥	٢	٣٩,٣٣٤	١١,٨٠١
الخطأ (بالطرح)	١٣,٣٣٢٥	٤	٣,٣٣٣	
المجموع الكلي	٩٢	٦		

وبما أن إحصائية الاختبار (١١,٨٠١) أكبر من القيمة الحرجة (المجدولة) بمستوى معنوية ٩٥و٠ ، فلا يمكن قبول فرضية العدم •

(٢) $ر^٢ = \frac{م م ر}{م م ك}$ (٥٥)

$= \frac{٧٨,٦٦٧٥}{٩٢}$

= ٠,٨٥٥

(٣) باستخـــدام جدول تحليل التبـــاين الســـابق يتضـــح أن القيمـة التقـديرية للتبـاين $ع^{٢}$ = ٣,٣٣٣ ومصفوفة التشتت هى :

$$ع^{٢}\,(\underline{س}'\,\underline{س})^{-١} = \frac{٣,٣٣٣}{١١٦٢٠٠}\begin{bmatrix} ٧٥٦٠- & ٢٨٧٠٠- & ٨٨٣٧٦٠ \\ ١٤٠- & ١١٩٠ & ٢٨٧٠٠- \\ ٧٠٠ & ١٤٠- & ٧٥٦٠- \end{bmatrix}$$

(٤) من مصفوفة التشتت السابقة:

$$ع_{أ}^{٢} = \frac{٣,٣٣٣}{١١٦٢٠٠} \times ٨٨٣٧٦٠$$

$$= ٢٥,٣٤٩$$

$$ع_{أ} = ٥,٠٣٥$$

$$ع_{ب}^{٢} = \frac{١١٩٠}{١١٦٢٠٠} \times ٣,٣٣٣$$

$$= ٠و٠٣٤$$

$$ع_{ب} = ٠,١٨$$

$$ع_{جـ}^{٢} = \frac{٧٠٠}{١١٦٢٠٠} \times ٣,٣٣٣$$

$$= ٠,٠٢$$

$$ع_{جـ} = ٠,١٤$$

إحصائية الاختبار هى :

$$\frac{أ}{ع_{أ}} = \frac{١٥,٤٦٩-}{٥,٠٣٥}$$

$$= ٣,٠٧-$$

$$\frac{ب}{ع_{ب}} = \frac{٠,٦٧٦٥}{٠,١٨}$$

$= ٣,٧٦$

$$\frac{جـ}{ع_{جـ}} = \frac{٠,٣٦٧٥}{٠,١٤}$$

$= ٢,٦٣$

وبما أن إحصائية الاختبار الخاصة بكل متغير أكبر من ٢ ، فلا يمكن قبول أى من الفرضيات التالية :

$ف_{٠}$: أ = صفراً

أو :

$ف_{٠}$: ب = صفراً

أو :

$ف_{٠}$: جـ = صفراً

تطبيق المصفوفات على النموذج المركزى :

يفضل الكثيرون استخدام النموذج المركزى ؛ لأن إيجاد مقلوب المصفوفة (سَ سَ) يكون أسهل من سابقه فى حالة تنفيذه يدوياً . وتجدر الإشارة هنا إلى أن النموذج المركزى يكون على النحو الآتى :

$صَ_{ر}$ = أ + ب ($س_{١ر}$ – $\bar{س}_{١}$) + جـ ($س_{٢ر}$ – $\bar{س}_{٢}$) + ... (٣٨)

= أ + ب $س_{١ر}$ + جـ $س_{٢ر}$ + ...

حيث $س_{١ر}$ ، $س_{٢ر}$ ، ...... هى انحرافات المتغيرات عن أوساطها .

مثال (٧ , ١١) :

استخدم بيانات المثال (٣) ، أو المثال (٦) ، لإيجاد معادلة الانحدار مستخدماً النموذج المركزى، والبيانات هى :

ص	س ١	س ٢
٣	٢٣	٩
٦	٢٨	١٢
٥	٢١	١٢
٨	٢٣	٢٢
١٠	٣٠	١٦
١٥	٣٢	١٨
٩	٣٥	٢٣
٥٦	١٨٢	١١٢

الحل :

$$\bar{س}_{١} = \frac{١٨٢}{٧}$$

$$= ٢٦$$

$$\bar{س}_{٢} = \frac{١١٢}{٧}$$

$$= ١٦$$

وبالتالى فمصفوفة الانحرافات للنموذج المركزى هى :

$$\underline{س} = \begin{pmatrix} ١ & -٣ & -٧ \\ ١ & ٢ & -٤ \\ ١ & -٥ & -٤ \\ ١ & -٣ & ٦ \\ ١ & ٤ & ٠ \\ ١ & ٦ & ٢ \\ ١ & -١ & ٧ \end{pmatrix}$$

و بالتالى فإن :

$$(\underline{\text{سَ}}\ \underline{\text{س}})^{-1} = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 100 & 20 \\ 0 & 20 & 170 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{1190}{116200} & \frac{-140}{116200} \\ 0 & \frac{-140}{116200} & \frac{700}{116200} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{17}{1660} & \frac{-2}{1660} \\ 0 & \frac{-2}{1660} & \frac{10}{1660} \end{bmatrix}$$

$$\underline{\text{سَ}}\ \underline{\text{ص}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -3 & 2 & -5 & -3 & 4 & 6 & -1 \\ -7 & -4 & -4 & 6 & 0 & 2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 6 \\ 5 \\ 8 \\ 10 \\ 15 \\ 9 \end{bmatrix}$$

$$= \begin{bmatrix} 56 \\ 75 \\ 76 \end{bmatrix}$$

$$\underline{\text{ب}} = (\underline{\text{سَ}}\ \underline{\text{س}})^{-1}\ \underline{\text{سَ}}\ \underline{\text{ص}} = \begin{bmatrix} \frac{1}{7} & 0 & 0 \\ 0 & \frac{17}{1660} & \frac{-2}{1660} \\ 0 & \frac{-2}{1660} & \frac{10}{1660} \end{bmatrix} \begin{bmatrix} 56 \\ 75 \\ 76 \end{bmatrix}$$

$$= \begin{bmatrix} ٨ \\ \frac{١١٢٣}{١٦٦٠} \\ \frac{٦١٠}{١٦٦٠} \end{bmatrix}$$

إذاً :

$\text{أ} = ٨ = \bar{\text{ص}}$

$\text{ب} = \frac{١١٢٣}{١٦٦٠}$

$\text{جـ} = \frac{٦١٠}{١٦٦٠}$

$$\therefore \hat{\text{ص}}_{\text{ر}} = ٨ + \frac{١١٢٣}{١٦٦٠}(\text{س}_{١\text{ر}} - ٢٦) + \frac{٦١٠}{١٦٦٠}(\text{س}_{٢\text{ر}} - ١٦)$$

$$= ٨ - \frac{١١٢٣}{١٦٦٠} \times ٢٦ - \frac{٦١٠}{١٦٦٠} \times ١٦ + \frac{١١٢٣}{١٦٦٠}\text{س}_{١\text{ر}} + \frac{٦١٠}{١٦٦٠}\text{س}_{٢\text{ر}}$$

$$= -١٥,٤٦٩ + ٠,٦٧٦٥\ \text{س}_{١\text{ر}} + ٣٦٧٥\ \text{س}_{٢\text{ر}}$$

وهو نفس نموذج المثال (٣) ونموذج المثال (٦) .

أما بالنسبة لجدول تحليل التباين فهو يتكون من :

المجموع الكلى (م م ك) = $\underline{\text{ص}'}\ \underline{\text{ص}} - \text{ن}\ \bar{\text{ص}}^{٢}$.

$= \sum \text{ص}^{٢}_{\text{ر}}$

$= ٥٤٠$

$$\underline{\text{ب}'}\ \underline{\text{س}'}\ \underline{\text{ص}} = \begin{pmatrix} ٨ & \frac{١١٢٣}{١٦٦٠} & \frac{٦١٠}{١٦٦٠} \end{pmatrix} \begin{pmatrix} ٥٦ \\ ٧٥ \\ ٧٦ \end{pmatrix}$$

$= ٤٤٨ + ٧٨,٦٧$

$= ٥٢٦,٦٧$

وهو نفس المقدار الوارد فى المثال (٦) . إذاً فجدول تحليل التباين يتم بنفس الأسلوب السابق ، كذلك يلاحظ من المصفوفة (سَ/ سَ)$^{-1}$ فى المثال السابق والمثال رقم (٦) أن مصفوفة التشتت لم تتغير .

البرنامج التالى يقوم بإيجاد معادلة الانحدار الخطى البسيط وجدول تحليل التباين باستخدام المصفوفات . استخدمنا فى هذا البرنامج التعليمات الخاصة بالمصفوفات مثل :

MAT READ
TRN
INV

وتعليمات الضرب والجمع وغيرهما ، وكلها تحدثنا عنها فى معرض حديثنا عن تعليمات لغة بيسك فى الفصل الثانى .

هذا البرنامج يعالج مصفوفات بأبعاد مختلفة لاستخدامه لمصفوفة بأبعاد مختلفة عن تلك التى فى البرنامج فينبغى فقط تعديل عبارات (DIM) وكذلك العبارة (720).

أما إذا كانت نسخة بيسك التى لديك لا تحتوى على تعليمات المصفوفات (MAT)، فيمكنك استخدام البرنامجين اللذين بعده ، حيث يقوم أحدهما بإجراء العمليات الأساسية للمصفوفات، والثانى يقوم بإيجاد معكوس المصفوفة (INV)، والبرنامجان لا يستخدمان تعليمات المصفوفات .

```
10  REM   برنامج لايجاد معادلة الانحدار الخطى البسيط
15  REM   وجدول تحليل التباين باستخدام المصفوفات
20  DIM X(7,2),Y(7,1),A(2,7),B(2,2),C(2,2),V(2,2),D(2,1),E(2,1),Z(7,1)
30  DIM P(1,1),T(1,1),H(1,1),R(1,1),S(1,1),G(1,1),Q(1,2),F(1,1),N(3,1)
31  DIM O(1,1),I(1,1),K(1,1),U(1,1),L(7,1),M(1,7),J(3,2),W(1,1)
35  N=2     REM  NO OF REGRESSION PARAMETERS
40  MAT READ X,Y
45  MAT READ J        REM VALUES TO BE USED FOR FORECAST
50  MAT A=TRN(X)      REM TRANSPOSE OF MATRIX X
60  MAT B=A*X         REM GIVES N SUMS & SUMS OF SQUARES & CROSS-PRODUCTS
70  MAT C=INV(B)      REM USED FOR PARAMETERS AND VARIANCE COVARIANCE
80  MAT D=A*Y         REM GIVES SUMS AND CROSS-PRODUCTS WITH Y
90  MAT E=C*D         REM COLUMN VECTOR OF PARAMETERS
100 G(1,1)=(D(1,1)/B(1,1))**2*B(1,1)
130 MAT Q=TRN(E)
140 MAT P=Q*D
150 MAT P=P - G
160 A1=(B(1,1)-1)**(-1)
180 N1=(N-1)**(-1)
190 MAT W=(N1)*P
220 MAT M=TRN(Y)
230 MAT T=M*Y
240 MAT T=T-G
250 MAT U=INV(T)
260 K(1,1)=B(1,1) - N
265 MAT I=INV(K)
270 MAT H=T - P
280 MAT S=H*I
```

```
290 MAT O=INV(S)
300 MAT F=O*W
330 MAT R=P*U
340 E1=S(1,1)
350 MAT V=(E1)*C       REM VARIANCE COVARIANVE MATRIX
360 S1=SQR(V(1,1))     REM STANDARD DEV. OF FIRST PARAMETER
370 S2=SQR(V(2,2))     REM STANDARD DEV. OF SECOND PARAMETER
380 T1=E(1,1)/S1       REM T VALUE OF FIRST PARAMETER
390 T2=E(2,1)/S2       REM T VALUE OF SECOND PARAMETER
400 MAT Z=X*E          REM PREDICTED VALUES
410 MAT L=Z-Y          REM RESIDUAL
420 MAT N=J*E          REM FORECAST
490 PRINT , E(1,1); 'تقدير المعلم ا ='
500 PRINT , E(2,1); 'تقدير المعلم ب ='
520 PRINT ,R(1,1); 'معامل التحديد ='
525 PRINT
527 PRINT
530 PRINT USING 810
540 PRINT USING 820
550 PRINT USING 830
560 PRINT USING 840
570 PRINT USING 850
580 PRINT USING 860,F(1,1),W(1,1),N-1,P(1,1)
585 PRINT
590 PRINT USING 870,S(1,1),K(1,1),H(1,1)
600 PRINT USING 880
610 PRINT USING 890,B(1,1)-1,T(1,1)
615 PRINT
618 PRINT
620 PRINT 'مصفوفه التشتت'
630 MAT PRINT V
635 PRINT
637 PRINT
640 PRINT ,S1;'(ا) الانحراف المعيارى للمعلم'
650 PRINT
660 PRINT ,S2;'(ب) الانحراف المعيارى للمعلم'
665 PRINT
666 PRINT
670 PRINT ,T1;'(ا) قيمه ت للمعلم'
680 PRINT
690 PRINT ,T2;'(ب) قيمه ت للمعلم'
700 PRINT
710 PRINT
715 PRINT USING 900
716 PRINT USING 920
717 PRINT USING 930
718 PRINT USING 950
720 FOR I=1 TO 7
730 PRINT USING 960,X(I,2),Y(I,1),Z(I,1),L(I,1)
735 PRINT
740 NEXT I
741 PRINT USING 970
742 PRINT USING 980
743 PRINT USING 990
744 PRINT USING 1000
750 FOR I=1 TO 3
760 PRINT USING 1010,J(I,2),N(I,1)
770 NEXT I
810 :                      جدول تحليل التباين
820 :   ------------------------------------------------------------------
830 :   ف ٥,١        متوسط        درجات       مجموع           مصدر
840 :                المربعات     الحريه      المربعات        التباين
850 :   ------------------------------------------------------------------
860 :  ####.####    ####.###      ##       #####.####      الانحدار (ب)
870 :               ####.###      ##       #####.####      الخطأ
880 :               ---------------------------------------------------
890 :                             ##      ######.####      المجموع الكلى
900 :                   جدول التقدير باستخدام المعادله
920 :   ------------------------------------------------------------------
930 :               X          Y          التقدير          خطأ التقدير
950 :   ------------------------------------------------------------------
960 :            ####.##    ####.##       ####.####        ###.####
970 :          جدول التنبوء باستخدام المعادله
980 :   ------------------------------------------------------------------
990 :               X        التنبوء
1000:   ------------------------------------------------------------------
1010:            ####.##    ####.##
1020 DATA 1,10,1,15,1,14,1,16,1,11,1,20,1,19
1030 DATA 25,35,30,40,24,50,48
1040 DATA 1,17,1,18,1,25
9999 END
```

المخرجات

تقدير المعلم ا = -5.25
تقدير المعلم ب = 2.749985
معامل التحديد = .9653304

جدول تحليل التباين

مصدر التباين	مجموع المربعات	درجات الحريه	متوسط المربعات	ف ١،٥
الانحدار (ب)	635.1875	1	635.187	139.2192
الخطا .	22.8125	5	4.562	
المجموع الكلى	658.0000	6		

مصفوفه التشتت

12.87273	-.8147299
-.8147299	5.431531E-02

الانحراف المعيارى للمعلم (ا) 3.587858
الانحراف المعيارى للمعلم (ب) .2330564

قيمه ت للمعلم (ا) -1.463268
قيمه ت للمعلم (ب) 11.79965

جدول التقدير باستخدام المعادله

X	Y	التقدير	خطا التقدير
10.00	25.00	22.2498	-2.7502
15.00	35.00	35.9998	0.9998
14.00	30.00	33.2498	3.2498
16.00	40.00	38.7498	-1.2502
11.00	24.00	24.9998	0.9998
20.00	50.00	49.7497	-0.2503
19.00	48.00	46.9997	-1.0003

جدول التنبوء باستخدام المعادله

X	التنبوء
17.00	41.50
18.00	44.25
25.00	63.50

الشروط الواجب توفرها فى معادلة الانحدار الخطى :

هناك عدة شروط يجب توفرها فى البيانات والنموذج ، وبدونها لا يكون النموذج المستخرج عملياً فى الوصف أو التقدير . ويتضح ذلك من فشل إحصائيات المعالم في اجتياز القيمة الحرجة (حوالى اثنتين) ، أو فى ضعف معامل التحديد، وهذه ظواهر كثيرة الحدوث أثناء التطبيقات العملية خاصة فى المجال الاقتصادى والمجال الاجتماعى . أما أهم هذه الشروط فيمكن إيجازها فيما يلى :

أ ـ اختيار النموذج المناسب ويتضمن ذلك :

١ ـ أن تكون العلاقة بين المتغيرات خطية .

٢ ـ ألا يكون هناك متغير ذو علاقة قد تم استبعاده ، أو متغير ليست له علاقة أضيف للنموذج . ومن المعروف أن زيادة عدد المتغيرات تزيد من قيمة معامل التحديد؛ لأنها تضعف عدد درجات حرية الخطأ دون مبرر ، فيبدو معامل التحديد كبيراً مع ضعف فى معالم النموذج .

ب ـ عدم وجود أخطاء فى القياس أثناء جمع البيانات .

جـ ـ عدم وجود ارتباط ذاتى (Autocorrelation) بين المتغيرات .

د - يجب أن تكون الأخطاء موزعة توزيعاً طبيعياً ، كما يجب ألا يكون بينها ارتباط ، وألا تتأثر طردياً أو عكسياً بقيم المتغيرات (Homoskedasticity) ، كذلك يجب أن يساوى وسطها الحسابى صفراً . وللتأكد من توفر هذا الشرط يتم تنفيذ رسم بيانى للأخطاء على القيم التقديرية من النموذج .

يتضح من ذلك أنه من المتوقع فى حالات كثيرة ألا تحقق إحصائيات الاختبار للمعالم القيم المطلوبة لتصبح ذات فاعلية . وبالنظر إلى المعادلة رقم (٢٨) أو المعادلة رقم (٢٩) يتضح أن تباين المعلم يتناسب تناسباً عكسياً مع تباين المتغير المستقل ($س_ر$) .فإذا كانت قيم المتغير متقاربة مع بعضها أصبح تباين المعلم كبيراً، بيد أن إحصائية الاختبار لكل معلم تتناسب تناسباً عكسياً مع تباينه . إذاً فربما يعزى الفشل فى ضعف قيمة إحصائية الاختبار إلى تجانس القيم العينية للمتغير المستقل ، وهو أمر لا يمكن معالجته إلا بزيادة عدد المتغيرات ، وذلك بإضافة قيم عينية أخرى أكثر تطرفاً .

أما ارتكاب الخطأ من النوع الثانى أثناء اختبار فرضية العدم لمعالم النموذج ، أو عدم اختيار النموذج المناسب ، فربما يؤدى إلى نفس النتيجة ، أو ضعف معامل التحديد . وفى هذه الحالة لا بد من تحويل البيانات التى اعتمد عليها النموذج (TRANSFORMATION) .

كذلك قد يتسبب الارتباط الشديد بين المتغيرات المستقلة (Multicollinearity) في إضعاف القيم الإحصائية للمعالم ، مع وجود قيمة عالية لمعامل التحديد، فيصبح النموذج غير صالح للوصف، وبمعالم ذات اتجاهات تخالف المنطق والواقع في أكثر الحالات . وفى هذه الحالة لا يمكن الاعتماد على النموذج ، إلا إذا أُضعِف ذلك الارتباط الشديد . وهناك طرق عديدة لإضعافه ، أهمها زيادة حجم العينة ، أو إلغاء المتغير الأكثر ارتباطاً (أو دمجه في متغير آخر) بالمتغيرات الأخرى ويمكن الكشف عن ذلك المتغير بإعداد نماذج خطية للمتغيرات المستقلة بعضها عن بعض ، وتحديد النموذج الذى يتميز باعلى معامل للتحديد .

فيما يلى برنامج لإجراء العمليات الأولية على المصفوفات بدون استخدام تعليمات المصفوفات (MAT).

```
10  REM   برنامج لاجراء العمليات الحسابيه الاوليه للمصفوفات
20  REM   وبدون استخدام تعليمات المصفوفات  MAT
25  DIM X(7,3),Y(7,1),T(3,7),E(3,3),H(3,1)
30  REM DIM X(R,Z),Y(Z,Q),T(Z,R),E(Z,Z),H(Z,Q)
35  REM X IS THE MATRIX OF INDEP.
40  REM Y IS THE MATRIX OF DEP.
60  REM T IS THE TRANSPOSE OF X.
70  REM E IS THE MATRIX T*X.
80  REM H IS THE MATRIX T*Y.
90  REM TRANSPOSITION PROC. NEXT
100  READ R,Z,Q
110  FOR I=1 TO R
120   FOR J=1 TO Z
130    READ X(I,J)
140   NEXT J
150  NEXT I
160  FOR I=1 TO R
165   FOR J=1 TO Q
170    READ Y(I,J)
180   NEXT J
185  NEXT I
190  FOR I = 1 TO Z
200   FOR J = 1 TO R
210    T(I,J) = X(J,I)
220   NEXT J
230  NEXT I
235  FOR K = 1 TO Z
240   FOR J = 1 TO Z
250    E(K,J) = 0
260    FOR I = 1 TO R
270     E(K,J) = E(K,J) + T(K,I) * X(I,J)
280    NEXT I
290   NEXT J
```

```
300  NEXT K
310  FOR K = 1 TO Z
325  FOR J = 1 TO Q
330  H(K,J) = 0
340  FOR I = 1 TO R
350  H(K,J) = H(K,J) + T(K,I) * Y(I,J)
360  NEXT I
370  NEXT J
380  NEXT K
390 PRINT ' MATRIX (X)'
400  FOR I=1 TO R
410   FOR J=1 TO Z
420    PRINT X(I,J),
430   NEXT J
440   PRINT
450  NEXT I
455   PRINT
460 PRINT 'MATRIX (T)'
470  FOR I=1 TO Z
480   FOR J=1 TO R
490    PRINT T(I,J),
500   NEXT J
510   PRINT
520  NEXT I
525   PRINT
530 PRINT 'MATRIX (E)'
540  FOR I=1 TO Z
550   FOR J=1 TO Z
560    PRINT E(I,J),
570   NEXT J
580   PRINT
590  NEXT I
595   PRINT
600 PRINT 'MATRIX (Y)'
610  FOR I=1 TO R
620   FOR J=1 TO Q
630    PRINT Y(I,J),
640   NEXT J
650   PRINT
660  NEXT I
665   PRINT
670  PRINT 'MATRIX (H)'
680  FOR I = 1 TO Z
690   FOR J = 1 TO Q
700    PRINT H(I,J),
710   NEXT J
720   PRINT
730  NEXT I
740  PRINT
750 DATA 7,3,1
760 DATA 1,23,9,1,28,12,1,21,12,1,23,22,1,30,16,1,32,18,1,25,23,3,6,5,8
770 DATA 10,15,9
999 END
```

المخرجات

```
MATRIX (X)
1    23    9
1    28    12
1    21    12
1    23    22
1    30    16
1    32    18
1    25    23
MATRIX (T)
1
1    1     1     1     1
23   28    21    23    30
25
9    12    12    22    16
23
MATRIX (E)
7    182   112
182  4812  2932
112  2932  1962
MATRIX (Y)
3
6
5
8
10
15
9
MATRIX (H)
56
1531
972
```

البرنامج التالى يقوم بإيجاد مقلوب مصفوفة ذات أبعاد 5x5 . لاستخدام البرنامج لمصفوفة ذات أبعاد مختلفة يلزمك فقط التعديلات التالية :

- عبارات (DIM) فى السطر (20) .
- القيمة فى عبارة (DATA) فى السطر (40) حيث توضع بعد المصفوفة الجديدة .

يقوم البرنامج كذلك باختيار ما إذا كانت المصفوفة وحيدة (SINGULAR) وفى هذه الحالة فإنه يعطى رسالة بذلك.

```
10  REM    MAT برنامج لايجاد مقلوب المصفوفه بدون استخدام تعليمات
20  DIM X(5,5),E(5,10),H(5,5)
30  READ R       REM NO. OF ROWS
40  DATA 5
60  C=2*R        REM NO. OF COLUMNS
70   FOR I=1 TO R
80    FOR J=1 TO C/2
90     READ X(I,J)
100   NEXT J
110  NEXT I
120  FOR I=1 TO R
130   FOR J=1 TO C/2
140    E(I,J)=X(I,J)
150   NEXT J
160  NEXT I
170  FOR I=1 TO R
180   FOR J=R+1 TO C
190    E(I,J)=0
200    IF J<>I+R THEN 220
210    E(I,J)=-1
220   NEXT J
230  NEXT I
330 FOR M=1 TO R
340  FOR K=1 TO R
345   FOR J=1+M TO C
350    IF K= M THEN 410
355    IF M > 1 THEN 362
360    P = 1
361    GO TO 363
362    P= E(M-1,M-1)
363    FOR T=1 TO R
365     IF E(T,T) < > 0 THEN 405
370     FOR I = T+1 TO R
375      IF E(I,T) < > 0 THEN 390
380     NEXT I
385     PRINT ' A ZERO PIVOT.  HENCE, SINGULAR MATRIX'
386     GO TO 999
390     FOR L = 1 TO C
391      B = E(I,L)
392      E(I,L) = E(T,L)
393      E(T,L) = B
394     NEXT L
395    NEXT T
405    E(K,J)=(E(M,M)*E(K,J)-E(M,J)*E(K,M))/P
410   NEXT J
415  NEXT K
418  FOR K=1 TO R
419  FOR J=1+M TO C
420    IF K<>M THEN 427
423     E(M,J)= - E(M,J)
427   NEXT J
428  NEXT K
430 NEXT M
```

```
432 IF  E(R,R) <> 0 THEN 440
434 PRINT '2  SINGULAR MATRIX'
435 GO TO 999
440 FOR K=1 TO R
450  FOR J=R+1 TO C
460   E(K,J)=E(K,J)/E(R,R)
470   H(K,J-R)=E(K,J)
490  NEXT J
500 NEXT K
505 PRINT '          مقلوب المصفوفه'
507 PRINT '          ______________'
510 FOR I=1 TO R
520  FOR J=1 TO C/2
530   PRINT H(I,J),
540  NEXT J
550  PRINT
560 NEXT I
570 PRINT
580 PRINT 'DET =';E(R,R)
590 PRINT 'NOTE: SIGN OF THE DET. MIGHT BE CHANGED IN CASE'
600 PRINT '      OF CHANGING TWO ADJACENT ROWS'
650 DATA 5,0,0,0,0,0,10,0,0,0,0,0,2,-3,4,0,0,7,-1,3,0,0,-1,2,-2
999  END
```

المخرجات

مقلوب المصفوفه

```
.2          0           0               0                 0
0           .1          0               0                 0
0           0           -.3636363       .1818181          -.4545454
0           0           1               0                 2
0           0           1.181818        -9.090906E-02     1.727272

DET = 550
NOTE: SIGN OF THE DET. MIGHT BE CHANGED IN CASE
      OF CHANGING TWO ADJACENT ROWS
```

تمارين

١ - إذا كانت :

$\sum$ س$_{ر}$ = ٣٧ ، $\sum$ س$_{ر}^{٢}$ = ١٦٠ ، $\sum$ س$_{ر}$ ص$_{ر}$ = ٦٢٥ ، $\sum$ ص$_{ر}$ = ١٥٦

$\sum$ ص$_{ر}^{٢}$ = ٥ر٢٥٤٥ ، ن = ١٠

فأوجد معادلة الانحدار الخطى للمتغير ص$_{ر}$ على المتغيرس$_{ر}$.

٢ - تأكد من أن خط الانحدار فى المعادلة الخاصة بالسؤال السابق يمر بنقطة الوسطين ، ثم قدر قيمة ص$_{ر}$ إذا كانت س$_{ر}$ = ١٠ ، ٩ ، ١٢ ، ٧

٣ - أوجد معامل التحديد لمعادلة السؤال الأول .

٤ - أوجد انحراف كل معلم من معالم الانحدار فى السؤال الأول ، ثم قدر فترات الثقة للقيمة التقديرية للمتغير التابع بمستوى معنوية ٥٪ إذا كانت :

س$_{ر}$ = ١٠ ، ٩ ، ١٢ ، ٧

٥ - البيانات التالية تمثل عينة من محطات البنزين وعدد الطلمبات لصب البنزين فى كل محطة وكانت البيانات كالآتى :

رقم المحطة	عدد الطلمبات	كمية البنزين في يوم واحد بآلاف اللترات التى تم بيعها
١	٣	٢,٧
٢	٦	٤
٣	١	٠,٣
٤	١١	١٢
٥	٧	٣
٦	٨	٦

أوجد معادلة انحدار المبيعات على عدد الطلمبات ، واختبر دقتها، ثم قدر الكمية المبيعة فى محطة بنزين بها ١٠ طلمبات ، ومحطة بها ٩ طلمبات، ومحطة ثالثة بها ٥ طلمبات .

٦ ـ أوجد حدود الثقة بمستوى معنوية ٥٪ للقيم التقديرية فى السؤال السابق ، وارسم ذلك بيانياً.

٧ ـ البيانات التالية تمثل عدد العاملين بإحدى المؤسسات خلال عشر سنوات .
أوجد معادلة الانحدار الخطى وقدر عدد العاملين فى عام ١٤١١هـ، والبيانات هى :

السنة	١٣٩٦	١٣٩٧	١٣٩٨	١٣٩٩	١٤٠٠	١٤٠١	١٤٠٢	١٤٠٣	١٤٠٤	١٤٠٥
عدد العاملين	٥٠	٦٠	٦٥	٧٥	٨٠	٩٠	١٠١	١٠٩	١١٠	١١٩

٨ - إذا كانت :

ص ر = ٤ , ٩٢-٣س ر

فأوجـد الارتباط بين المتغيرين إذا كان معامل التحديد ٩٠٪ ، والانحراف المعيارى للمتغير ص يساوى ٢٥ ، والوسط الحسابي للمتغير س يساوى ١٠ .
كذلك أوجد الانحراف المعيارى للمتغير س ، والوسط الحسابي للمتغير ص.

٩ - إذا كانت :

س ر = ١ ، ٢ ، ٣ ، ٤ ، ٥ ، ٦ ، ٧ ، ٨

ص ر = ٦ ، ١٠ ، ٦ ، ١٠ ، ٦ ، ١٠ ، ٦ ، ١٠

فأوجد معادلة انحدار ص على س واختبر دقة المعادلة .

١٠ ـ البيانات التالية تمثل تكلفة الصيانة بآلاف الريالات سنوياً، والعمر لعدد من الناقلات، والبيانات هى :

التكلفة	٧	٤	٤	٩	١	٨
العمر بالسنوات	٣	٥	٤	٧	٢	٦

أوجد معادلة انحدار التكلفة على العمر ، وأوجد معامل التحديد والانحراف المعيارى لكل معلم ، واختبر مقدرته على الوصف والتنبؤ . كذلك قدر تكلفة الصيانة لشاحنة عمرها ٨ سنوات ، وشاحنة أخرى عمرها عام واحد ، وشاحنـــة جديدة .

١١ - إذا كانت :

س = ٩ ، ١٠ ، ١١ ، ١٢ ، ١٣ ، ١٤ ، ١٥ ، ١٦ ، ١٧ ، ١٨ ، ١٩

ص = -١٠ ، -٨ ، -٦ ، -٤ ، -٢ ، ٠ ، ٢ ، ٤ ، ٦ ، ٨ ، ١٠

فأوجد انحدار س على ص ، وأوجد معامل التحديد، واختبر معنوية المعالم بمستوى ٥٪.

١٢ - إذا كانت :

س = ١٨ ، ٢٠ ، ٢٢ ، ٢٤ ، ٢٦ ، ٢٨ ، ٣٠ ، ٣٢ ، ٣٤ ، ٣٦ ، ٣٨

ص = ٠ ، ٢ ، ٤ ، ٦ ، ٨ ، ١٠ ، ١٢ ، ١٤ ، ١٦ ، ١٨ ، ٢٠

فأوجد انحدار س على ص ، وأوجد معامل التحديد ، واختبر معنوية المعالم ، وقارن النموذج بنموذج السؤال السابق .

١٣ ـ استخدم بيانات السؤال الحادى عشر والسؤال الثانى عشر لاستخراج انحدار ص على س فى كل حالة ، وقارن بين كل حالتين .

١٤ ـ البيانات التالية تمثل عدد سكان إحدى المدن بالآلاف خلال الفترة من ١٣٩٨هـ حتى ١٤٠٦هـ :

السنة	١٣٩٨	١٣٩٩	١٤٠٠	١٤٠١	١٤٠٢	١٤٠٣	١٤٠٤	١٤٠٥	١٤٠٦
عدد السكان بالآلاف	٣١	٣٤	٣٧	٤٣	٤٤	٤٩	٧١	٩٨	١٠٧

أوجـد معـادلـة الانحدار الخطى لتقدير عدد السكان ووأجد معامل التحديد ، واختبر معنويـة المعالم بمستـوى معنوية ٥٪ ، وقدر عدد السكان في عام ١٤١٠هـ .

١٥ ـ استخدم بيانات السؤال السابق وأوجد معادلة الانحدار للوغريتهات البيانات بدلاً عن البيانات نفسها ، واختبر دقة المعادلة ومعنوية المعالم بمستوى معنوية ٥٪ ، وقدر عدد السكان في عام ١٤١٠هـ ، وقارن بين النموذج الحالى ونموذج السؤال السابق .

١٦ - اذا كانت :

ص ر	س ١ر	س ٢ر
٣٠	٤٤	١٥
٢٨	٤٦	١٤
٢٨	٤٤	١٤
٢٨	٤٥	١٥
٢٩	٤٦	١٦
٢٤	٤١	١٤
٢٧	٤٣	١٣
٢٦	٤٣	١٢
٢٧	٤٤	١٤
٢٥	٤٥	١١

فأوجد معادلة انحدار ص على المتغيرين س ١ و س ٢ ، واختبر دقة المعادلة ومعنوية المعالم وقدر قيمة ص ر إذا كانت س ١ = ٤٧ ، س ٢ = ١٦ .

١٧ - البيانات التالية تمثل عينة من ٨ عجول وأعمارها وأوزانها عند بداية تناول عشب خاص لزيادة الوزن ، ومن ثم زيادة الوزن فى كل حالة .

أوجـد معادلة انحدار زيادة الوزن على العمر والوزن الأصلى ثم اختبر دقة المعادلة ومعنوية المعالم بمستوى معنوية ٥٪، والبيانات هى :

العمر بالشهور	الوزن بالكيلوجرام	زيادة الوزن بالكيلوجرام
٥	٣٤	٨
١١	٢٥	٧
٤	٥٥	٨
٩	٤١	١١
٧	٥٤	١٠
٨	٥٠	٥
١٢	٤٨	٤
١٠	٦٣	٥

١٨ ـ البيانـات التـالية تمثـل متـوسط تكلفـة الصيانـة بالريالات شهرياً، وعدد الساعات الأسبوعية التى تعملها كل ماكينة، وعمر الماكينة بالشهور، وعدد المرات التى تتعرض فيها للصيانة شهرياً.

أوجـد معـادلة انحدار تكلفة الصيانة على المتغيرات الأخرى، وأوجد الارتباط بين كل متغيرين، واختبر دقة المعادلة ومعنوية المعالم بمستوى ٥٪ . والبيانات هى :

تكلفة الصيانة	عدد الساعات	العمر	عدد مرات الصيانة
٦٤٠	٩٠	٨٠	٤٠
٦١٠	٧٥	٥٠	١٠
٤٩٠	٤٥	٣٠	٢٠
٢٥٠	٤٥	٥	٥٠
٤٠٠	٤٥	١٥	٣٠
٧٩٠	٦٠	٧٠	١٠
٨٤٠	٩٠	٧٠	١٠
١٢٠	٤٥	١٥	٤٠
٣٦٠	٤٥	٣٠	٣٠
٥٠٠	٤٥	٧٥	١٠
٦١٠	١٢٠	٥٠	١٠
٤٩٠	٩٠	٦٠	٢٠
٢٥٠	٦٠	١٠	٥٠
٦١٠	١٢٠	٣٥	٤٠
٣١٠	١٢٠	١٥	٤٠
٢٠٠	٦٠	٥	٥٠
٢٢٠	٧٥	٢٠	٥٠
٣٢٠	٩٠	٤٠	٤٠

١٩ - البيانـات التـالية تمثل بالشهور المدة التى عملها كل موظف بإحدى المؤسسات ذات القسمين الرجالى والنسوى قبل أن ينتقل الموظف إلى جهة عمل أخرى . وتمثل نفس البيانـات عمـر كل موظف عند التحاقه بالعمل ، والدرجات التى حصل عليها فى مسابقة الالتحاق بالوظيفة (كنسبة مئوية) . هذا ويمثل الصفر المرأة والواحد الرجل .

والمطلوب هو تقدير معادلة الانتظام فى الوظيفة على المتغيرات الأخرى مع اختبار دقتها للتنبؤ والوصف .

المدة بالشهور	العمر	الجنس	درجات المسابقة	ملحوظة
٤٠	٢٠	٠	٧٠	
١٠	٢٠	٠	١٠	
٢٠	٣٠	١	٧٠	
٩٠	٣٠	١	٢٠	
١٠	٢٠	١	٢٠	
٩٠	٤٠	١	٤٠	
صفر	٢٠	٠	١٠٠	انسحب بعد أقل من شهر .
١٠	٥٠	٠	٦٠	
٨٠	٢٠	١	٧٠	
١٠	٢٠	١	٨٠	
٢٠	٤٠	١	٥٠	
١٠	٤٠	١	٦٠	
٣٠	٣٠	١	٤٠	
٣٠	٥٠	١	٧٠	
٣٠	٢٠	١	٦٠	
١٢٠	٣٠	٠	٩	
٤٣٠	٢٠	٠	١٠٠	
٢٠	٣٠	١	٢٠	
١٠٠	٣٠	٠	٥٠	
٤٨٠	٢٠	١	٩٠	

٢٠ - إذا كانت :

ص =	١٠	٣	٥٠	٦٠	٨٠	١٠٠
س$_1$ =	٣	٨	١٠	١١	١٥	١٩
س$_2$ =	٧	١٧	٢١	٢٣	٣١	٣٩

فأوجد معادلة انحدار ص على س$_1$ و س$_2$ ، واختبر قدرتها على التنبؤ والوصف . ما هى التعديلات التى يجب أن تطرأ على المعادلة لتحسين مقدرتها على الوصف ؟

٢١ - اكتب برنامج بيسك لإيجاد معادلة الانحدار الخطى للبيانات بالسؤال (١) ، ومن ثم معامل التحديد .

٢٢ - اكتب برنامج بيسك لإيجاد معادلة الانحدار الخطى للبيانات بالسؤال (٥) .

٢٣ - مستخدماً البيانات بالسؤال (٧) اكتب برنامج بيسك لإيجاد معادلة الانحدار الخطى .

٢٤ - استخـدم البيانـات بالسؤال (١٠) فى برنامج بيسك لإيجاد معادلة الانحدار ومعامل التحديد والانحراف المعيارى .

٢٥ - استخدم البيانات فى السؤال (١١) فى برنامج وأوجد معامل التحديد .

الملاحق

- جدول (١) التوزيع الطبيعي
- جدول (٢) توزيع ت
- جدول (٣) توزيع مربع كاي
- جدول (٤) توزيع ف
- جدول (٥) التوزيع ذو الحدين
- جدول (٦) اختبار حسن المطابقة لعينتين صغيرتين متساويتين
- جدول (٧) اختبار حسن المطابقة لعينتين كبيرتين
- جدول (٨) اختبار فروق الرتب للأزواج المتقارنة
- جدول (٩) اختبار مجموع الرتب لعينتين
- جدول (١٠) قيم ر الحرجة لاختبار ز = صفرا
- جدول (١١) تحويل ر إلى ى
- جدول (١٢) دوال لغة بيسك في جهاز (IBM)

جدول (١)

التوزيع الطبيعى (ى)

t	0	1	2	3	4	5	6	7	8	9
−3.	0013									
−2.9	0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	0107	.0104	.0102	.0099	.0096	.0094	0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0227	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0300	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1921	.1894	.1867
−.7	.2420	.2389	.2358	.2326	.2297	.2266	.2236	.2206	.2177	.2148
−.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

المصدر :

LARSON (H.J.) , Introd . To Prob –Theory and Statistical Inference; wiley 1976, page 398.

تابع جدول(١)

t	0	1	2	3	4	5	6	7	8	9
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8079	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9700	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9773	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.	.9987									

جدول (٢)

توزيع ت

د	درجات الحرية = د ١ - أ .60	.75	.90	.95	.975	.99	.995	.9995
1	.325	1.000	3.078	6.314	12.706	31.821	63.657	636.619
2	.289	.816	1.886	2.920	4.303	6.965	9.925	31.598
3	.277	.765	1.638	2.353	3.182	4.541	5.841	12.941
4	.271	.741	1.533	2.132	2.776	3.747	4.604	8.610
5	.267	.727	1.476	2.015	2.571	3.365	4.032	6.859
6	.265	.718	1.440	1.943	2.447	3.143	3.707	5.959
7	.263	.711	1.415	1.895	2.365	2.998	3.499	5.405
8	.262	.706	1.397	1.860	2.306	2.896	3.355	5.041
9	.261	.703	1.383	1.833	2.262	2.821	3.250	4.781
10	.260	.700	1.372	1.812	2.228	2.764	3.169	4.587
11	.260	.697	1.363	1.796	2.201	2.718	3.106	4.437
12	.259	.695	1.356	1.782	2.179	2.681	3.055	4.318
13	.259	.694	1.350	1.771	2.160	2.650	3.012	4.221
14	.258	.692	1.345	1.761	2.145	2.624	2.977	4.140
15	.258	.691	1.341	1.753	2.131	2.602	2.947	4.073
16	.258	.690	1.337	1.746	2.120	2.583	2.921	4.015
17	.257	.689	1.333	1.740	2.110	2.567	2.898	3.965
18	.257	.688	1.330	1.734	2.101	2.552	2.878	3.922
19	.257	.688	1.328	1.729	2.093	2.539	2.861	3.883
20	.257	.687	1.325	1.725	2.086	2.528	2.845	3.850
21	.257	.686	1.323	1.721	2.080	2.518	2.831	3.819
22	.256	.686	1.321	1.717	2.074	2.508	2.819	3.792
23	.256	.685	1.319	1.714	2.069	2.500	2.807	3.767
24	.256	.685	1.318	1.711	2.064	2.492	2.797	3.745
25	.256	.684	1.316	1.708	2.060	2.485	2.787	3.725
26	.256	.684	1.315	1.706	2.056	2.479	2.779	3.707
27	.256	.684	1.314	1.703	2.052	2.473	2.771	3.690
28	.256	.683	1.313	1.701	2.048	2.467	2.763	3.674
29	.256	.683	1.311	1.699	2.045	2.462	2.756	3.659
30	.256	.683	1.310	1.697	2.042	2.457	2.750	3.646
40	.255	.681	1.303	1.684	2.021	2.423	2.704	3.551
60	.254	.679	1.296	1.671	2.000	2.390	2.660	3.460
120	.254	.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	.253	.674	1.282	1.645	1.960	2.326	2.576	3.291

المصدر: نفس المصدر السابق صفحة (٤٠٢)

جدول (٣)

التوزيع مربع كاى (ك ٢)

درجات الحرية	0.995	0.990	0.975	0.950	0.900	0.100	0.050	0.025	0.010	0.005
1	0.0^{4}393	0 0^{3}157	0 0^{3}982	0 0^{2}393	0 0158	2 71	3 84	5 02	6.63	7 88
2	0 0100	0 0201	0 0506	0 103	0 211	4 61	5 99	7 38	9.21	10.60
3	0 072	0 115	0 216	0 352	0 584	6 25	7.81	9 35	11.34	12 84
4	0 207	0 297	0 484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0 554	0 831	1 145	1.61	9.24	11.07	12.83	15 09	16 75
6	0 676	0 872	1 24	1.64	2 20	10 64	12.59	14 45	16 81	18 55
7	0 989	1 24	1 69	2 17	2 83	12 02	14 07	16 01	18 48	20 28
8	1 34	1 65	2 18	2 73	3 49	13 36	15 51	17 53	20.09	21 96
9	1 73	2 09	2 70	3 33	4.17	14 68	16.92	19 02	21.67	23 59
10	2.16	2 56	3 25	3 94	4.87	15 99	18 31	20 48	23 21	25 19
11	2 60	3 05	3 82	4 57	5 58	17 28	19 68	21 92	24 72	26 76
12	3 07	3 57	4 40	5 23	6 30	18 55	21 03	23 31	26 22	28 30
13	3.57	4 11	5 01	5 89	7 04	19 81	22 36	24 74	27 69	29 82
14	4 07	4.66	5 63	6 57	7 79	21 06	23 68	26 12	29.14	31 32
15	4 60	5 23	6 26	7 26	8 55	22 31	25 00	27 49	30 58	32 80
16	5 14	5 81	6 91	7 96	9 31	23 54	26.30	28 85	32 00	34 27
17	5 70	6 41	7 56	8 67	10 09	24 77	27.59	30 19	33 41	35 72
18	6 26	7 01	8 23	9 39	10 86	25 99	28 87	31 53	34 81	37 16
19	6 84	7 63	8 91	10 12	11 65	27 20	30 14	32 85	36 19	38 58
20	7 43	8 26	9 59	10 85	12 44	28 41	31 41	34 17	37.57	40 00
21	8 03	8 90	10 28	11 59	13 24	29 62	32 67	35 48	38 93	41 40
22	8 64	9 54	10 98	12 34	14 04	30 81	33 92	36 78	40 29	42 80
23	9 26	10 20	11 69	13 09	14 85	32.01	35 17	38 08	41 64	44 18
24	9 89	10 86	12 40	13 85	15.66	33 20	36 42	39 36	42 98	45.56
25	10 52	11 52	13 12	14 61	16 47	34 38	37 65	40 05	44 31	46 93
26	11 16	12 20	13 84	15 38	17 29	35 56	38 89	41 92	45 64	48 29
27	11 81	12 88	14 57	16 15	18 11	36 74	40 11	43 19	46 96	49 64
28	12 46	13 56	15 31	16 93	18 94	37 92	41 34	44 46	48 28	50 99
29	13 12	14 26	16 05	17 71	19.77	39 09	42.56	45.72	49 59	52 34
30	13 79	14 95	16 79	18 49	20 60	40 26	43 77	46 98	50 89	53 67
40	20 71	22 16	24 43	26 51	29 05	51 80	55 76	59 34	63 69	66 77
50	27 99	29 71	32 36	34 76	37 69	63 17	67 50	71 42	76 15	79 49
60	35 53	37 48	40 48	43 19	46 46	74 40	79 08	83 30	88 38	91 95
70	43 28	45 44	48 76	51 74	55 33	85 53	90 53	95 02	100 4	104 22
80	51 17	53 54	57 15	60 39	64 28	96 58	101 9	106 6	112 3	116 32
90	59 20	61 75	65 65	69 13	73 29	107 6	113 1	118 1	124 1	128 3
100	67 33	70 06	74 22	77 93	82 36	118 5	124 3	129 6	135 8	140 2
z_α	−2 58	−2 33	−1 96	−1 64	−1 28	+1 28	+1 64	+1 96	+2 33	+2 58

المصدر

Patchet (I.S.); Statistical Methods; Van Nostrand; New York; 1982 page 356.

جدول (٤) توزيع ف

مستوى المعنوية ٥٪

درجات حرية البسط

درجات حرية المقام	1	2	3	4	5	6	7	8	9	10	12	15	20	21	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

المصدر: Patchet (I.S.); Statistical Methods; Van Nostrand, NewYork, 1982, page 360

تابع جدول (٤)

توزيع ف مستوى المعنوية ١٪

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4,052	5,000	5,403	5,625	5,764	5,859	5,928	5,982	6,023	6,056	6,106	6,157	6,209	6,235	6,261	6,287	6,313	6,339	6,366
2	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.53	2.45	2.36	2.27	2.17
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

جدول (٥)

التوزيع ذو الحدين

ن	[r]	0.10	0.20	p 0.30	0.40	0.50
2	0	.8100	.6400	.4900	.3600	.2500
	1	.9900	.9600	.9100	.8400	.7500
3	0	.7290	.5120	.3430	2160	.1250
	1	.9720	.8960	.7840	.6480	.5000
	2	.9990	.9920	.9730	.9360	.8750
4	0	.6561	.4096	.2401	.1296	.0625
	1	.9477	.8192	.6517	.4752	.3125
	2	.9963	.9728	.9163	.8208	.6875
	3	.9999	.9984	.9919	.9744	.9375
5	0	.5905	.3277	.1681	.0778	.0312
	1	.9185	.7373	.5282	.3370	.1875
	2	.9914	.9421	.8369	.6826	.5000
	3	.9995	.9933	.9692	.9130	.8125
	4	1.0000	.9997	.9976	.9898	.9688
6	0	.5314	.2621	.1176	.0467	.0156
	1	.8857	.6554	.4202	.2333	.1094
	2	.9842	.9011	.7443	.5443	.3438
	3	.9987	.9830	.9295	.8208	.6562
	4	.9999	.9984	.9891	.9590	.8906
	5	1.00 0	.9999	.9993	.9959	.9844
7	0	.4783	.2097	.0824	.0280	.0078
	1	.8503	.5767	.3294	.1586	.0625
	2	.9743	.8520	.6471	.4199	.2266
	3	.9973	.9667	.8740	.7102	.5000
	4	.9998	.9953	.9712	.9037	.7734
	5	1.0000	.9996	.9962	.9812	.9375
	6	1.0000	1.0000	.9998	.9984	.9922

المصدر : نفس المصدر للتوزيع الطبيعى صفحة (٣٩٠).

تابع جدول (٥)

ن	[r]	0.10	0.20	p 0.30	0.40	0.50
8	0	.4305	.1678	.0576	.0168	.0039
	1	.8131	.5033	.2553	.1064	.0352
	2	.9619	.7969	.5518	.3154	.1445
	3	.9950	.9437	.8059	.5941	.3633
	4	.9996	.9896	.9420	.8263	.6367
	5	1.0000	.9988	.9887	.9502	.8555
	6	1.0000	.9999	.9987	.9915	.9648
	7	1.0000	1.0000	.9999	.9993	.9961
9	0	.3874	.1342	.0404	.0101	.0020
	1	.7748	.4362	.1960	.0705	.0195
	2	.9470	.7382	.4628	.2318	.0898
	3	.9917	.9144	.7297	.4826	.2539
	4	.9991	.9804	.9012	.7334	.5000
	5	.9999	.9969	.9747	.9006	.7461
	6	1.0000	.9997	.9957	.9750	.9102
	7	1.0000	1.0000	.9996	.9962	.9805
	8	1.0000	1.0000	1.0000	.9997	.9980
10	0	.3487	.1074	.0282	.0060	.0010
	1	.7361	.3758	.1493	.0464	.0107
	2	.9298	.6778	.3828	.1673	.0547
	3	.9872	.8791	.6496	.3823	.1719
	4	.9984	.9672	.8497	.6331	.3770
	5	.9999	.9936	.9527	.8338	.6230
	6	1.0000	.9991	.9894	.9452	.8281
	7	1.0000	.9999	.9984	.9877	.9453
	8	1.0000	1.0000	.9999	.9983	.9893
	9	1.0000	1.0000	1.0000	.9999	.9990
11	0	.3138	.0859	.0198	.0036	.0005
	1	.6974	.3221	.1130	.0302	.0059
	2	.9104	.6174	.3127	.1189	.0327
	3	.9815	.8389	.5696	.2963	.1133
	4	.9972	.9496	.7897	.5328	.2744
	5	.9997	.9883	.9218	.7535	.5000
	6	1.0000	.9980	.9784	.9006	.7256
	7	1.0000	.9998	.9957	.9707	.8867
	8	1.0000	1.0000	.9994	.9941	.9673

تابع جدول (٥)

ن	[r]	0.10	0.20	p 0.30	0.40	0.50
	9	1.0000	1.0000	1.0000	.9993	.9941
	10	1.0000	1.0000	1.0000	1.0000	.9995
12	0	.2824	.0687	.0138	.0022	.0002
	1	.6590	.2749	.0850	.0196	.0032
	2	.8891	.5583	.2528	.0834	.0193
	3	.9744	.7946	.4925	.2253	.0730
	4	.9957	.9274	.7237	.4382	.1938
	5	.9995	.9806	.8822	.6652	.3872
	6	.9999	.9961	.9614	.8418	.6128
	7	1.0000	.9994	.9905	.9427	.8062
	8	1.0000	.9999	.9983	.9847	.9270
	9	1.0000	1.0000	.9998	.9972	.9807
	10	1.0000	1.0000	1.0000	.9997	.9968
	11	1.0000	1.0000	1.0000	1.0000	.9998
13	0	.2542	.0550	.0097	.0013	.0001
	1	.6213	.2336	.0637	.0126	.0017
	2	.8661	.5017	.2025	.0579	.0112
	3	.9658	.7473	.4206	.1686	.0461
	4	.9935	.9009	.6543	.3530	.1334
	5	.9991	.9700	.8346	.5744	.2905
	6	.9999	.9930	.9376	.7712	.5000
	7	1.0000	.9988	.9818	.9023	.7095
	8	1.0000	.9998	.9960	.9679	.8666
	9	1.0000	1.0000	.9993	.9922	.9539
	10	1.0000	1.0000	.9999	.9987	.9888
	11	1.0000	1.0000	1.0000	.9999	.9983
	12	1.0000	1.0000	1.0000	1.0000	.9999
14	0	.2288	.0440	.0068	.0008	.0001
	1	.5846	.1979	.0475	.0081	.0009
	2	.8416	.4481	.1608	.0398	.0065
	3	.9559	.6982	.3552	.1243	.0287
	4	.9908	.8702	.5842	.2793	.0898
	5	.9985	.9561	.7805	.4859	.2120
	6	.9998	.9884	.9067	.6925	.3953
	7	1.0000	.9976	.9685	.8499	.6047
	8	1.0000	.9996	.9917	.9417	.7880
	9	1.0000	1.0000	.9983	.9825	.9102

تابع جدول (٥)

ن	[/]	0.10	0.20	0.30	0.40	0.50
	10	1.0000	1.0000	.9998	.9961	.9713
	11	1.0000	1.0000	1.0000	.9994	.9935
	12	1.0000	1.0000	1.0000	.9999	.9991
	13	1.0000	1.0000	1.0000	1.0000	.9999
15	0	.2059	.0352	.0047	.0005	.0000
	1	.5490	.1671	.0353	.0052	.0005
	2	.8159	.3980	.1268	.0271	.0037
	3	.9444	.6482	.2969	.0905	.0176
	4	.9873	.8358	.5155	.2173	.0592
	5	.9978	.9389	.7216	.4032	.1509
	6	.9997	.9819	.8689	.6098	.3036
	7	1.0000	.9958	.9500	.7869	.5000
	8	1.0000	.9992	.9848	.9050	.6964
	9	1.0000	.9999	.9963	.9662	.8491
	10	1.0000	1.0000	.9993	.9907	.9408
	11	1.0000	1.0000	.9999	.9981	.9824
	12	1.0000	1.0000	1.0000	.9997	.9963
	13	1.0000	1.0000	1.0000	1.0000	.9995
	14	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	.1853	.0281	.0033	.0003	.0000
	1	.5147	.1407	.0261	.0033	.0003
	2	.7892	.3518	.0994	.0183	.0021
	3	.9316	.5981	.2459	.0651	.0106
	4	.9830	.7982	.4499	.1666	.0384
	5	.9967	.9183	.6598	.3288	.1051
	6	.9995	.9733	.8247	.5272	.2272
	7	.9999	.9930	.9256	.7161	.4018
	8	1.0000	.9985	.9743	.8577	.5982
	9	1.0000	.9998	.9929	.9417	.7728
	10	1.0000	1.0000	.9984	.9809	.8949
	11	1.0000	1.0000	.9997	.9951	.9616
	12	1.0000	1.0000	1.0000	.9991	.9894
	13	1.0000	1.0000	1.0000	.9999	.9979
	14	1.0000	1.0000	1.0000	1.0000	.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000

تابع جدول (٥)

ن	[r]	0.10	0.20	p 0.30	0.40	0.50
17	0	.1668	.0225	.0023	.0002	.0000
	1	.4818	.1182	.0193	.0021	.0001
	2	.7618	.3096	.0774	.0123	.0012
	3	.9174	.5489	.2019	.0464	.0064
	4	.9779	.7582	.3887	.1260	.0245
	5	.9953	.8943	.5968	.2639	.0717
	6	.9992	.9623	.7752	.4478	.1662
	7	.9999	.9891	.8954	.6405	.3145
	8	1.0000	.9974	.9597	.8011	.5000
	9	1.0000	.9995	.9873	.9081	.6855
	10	1.0000	.9999	.9968	.9652	.8338
	11	1.0000	1.0000	.9993	.9894	.9283
	12	1.0000	1.0000	.9999	.9975	.9755
	13	1.0000	1.0000	1.0000	.9995	.9936
	14	1.0000	1.0000	1.0000	.9999	.9988
	15	1.0000	1.0000	1.0000	1.0000	.9999
	16	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	.1501	.0180	.0016	.0001	.0000
	1	.4503	.0991	.0142	.0013	.0001
	2	.7338	.2713	.0600	.0082	.0007
	3	.9018	.5010	.1646	.0328	.0038
	4	.9718	.7164	.3327	.0942	.0154
	5	.9936	.8671	.5344	.2088	.0481
	6	.9988	.9487	.7217	.3743	.1189
	7	.9998	.9837	.8593	.5634	.2403
	8	1.0000	.9957	.9404	.7368	.4073
	9	1.0000	.9991	.9790	.8653	.5927
	10	1.0000	.9998	.9939	.9424	.7597
	11	1.0000	1.0000	.9986	.9797	.8811
	12	1.0000	1.0000	.9997	.9942	.9519
	13	1.0000	1.0000	1.0000	.9987	.9846
	14	1.0000	1.0000	1.0000	.9998	.9962
	15	1.0000	1.0000	1.0000	1.0000	.9993
	16	1.0000	1.0000	1.0000	1.0000	.9999
	17	1.0000	1.0000	1.0000	1.0000	1.0000

تابع جدول (٥)

ن	[r]	0.10	0.20	p 0.30	0.40	0.50
19	0	.1351	.0144	.0011	.0001	.0000
	1	.4203	.0829	.0104	.0008	.0000
	2	.7054	.2369	.0462	.0055	.0004
	3	.8850	.4551	.1332	.0230	.0022
	4	.9648	.6733	.2822	.0696	.0096
	5	.9914	.8369	.4739	.1629	.0318
	6	.9983	.9324	.6655	.3081	.0835
	7	.9997	.9767	.8180	.4878	.1796
	8	1.0000	.9933	.9161	.6675	.3238
	9	1.0000	.9984	.9674	.8139	.5000
	10	1.0000	.9997	.9895	.9115	.6762
	11	1.0000	1.0000	.9972	.9648	.8204
	12	1.0000	1.0000	.9994	.9884	.9165
	13	1.0000	1.0000	.9999	.9969	.9682
	14	1.0000	1.0000	1.0000	.9994	.9904
	15	1.0000	1.0000	1.0000	.9999	.9978
	16	1.0000	1.0000	1.0000	1.0000	.9996
	17	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	.1216	.0115	.0008	.0000	.0000
	1	.3917	.0692	.0076	.0005	.0000
	2	.6769	.2061	.0355	.0036	.0002
	3	.8670	.4114	.1071	.0160	.0013
	4	.9568	.6296	.2375	.0510	.0059
	5	.9887	.8042	.4164	.1256	.0207
	6	.9976	.9133	.6080	.2500	.0577
	7	.9996	.9679	.7723	.4159	.1316
	8	.9999	.9900	.8867	.5956	.2517
	9	1.0000	.9974	.9520	.7553	.4119
	10	1.0000	.9994	.9829	.8725	.5881
	11	1.0000	.9999	.9949	.9435	.7483
	12	1.0000	1.0000	.9987	.9790	.8684
	13	1.0000	1.0000	.9997	.9935	.9423
	14	1.0000	1.0000	1.0000	.9984	.9793
	15	1.0000	1.0000	1.0000	.9997	.9941
	16	1.0000	1.0000	1.0000	1.0000	.9987
	17	1.0000	1.0000	1.0000	1.0000	.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000

جدول (٦)

القيم الحرجة لاختبار عينتين

صغيرتين متساويتين (حسن المطابقة)

CRITICAL VALUES OF *K* IN THE KOLMOGOROV-SMIRNOV TWO-SAMPLE TEST (small samples)

N	One-tailed test*		Two-tailed test†	
	$\alpha = .05$	$\alpha = .01$	$\alpha = .05$	$\alpha = .01$
3	3	—	—	—
4	4	—	4	—
5	4	5	5	5
6	5	6	5	6
7	5	6	6	6
8	5	6	6	7
9	6	7	6	7
10	6	7	7	8
11	6	8	7	8
12	6	8	7	8
13	7	8	7	9
14	7	8	8	9
15	7	9	8	9
16	7	9	8	10
17	8	9	8	10
18	8	10	9	10
19	8	10	9	10
20	8	10	9	11
21	8	10	9	11
22	9	11	9	11
23	9	11	10	11
24	9	11	10	12
25	9	11	10	12
26	9	11	10	12
27	9	12	10	12
28	10	12	11	13
29	10	12	11	13
30	10	12	11	13
35	11	13	12	
40	11	14	13	

المصدر :

Mason (D.R.) ; Statistical Techniques in Business and Economics; Third Edition , 1974 , IRWIN , Homewood page (637).

جدول (٧)

اختبار حسن المطابقة لعينتين كبيرتين

CRITICAL VALUES OF *D* IN THE KOLMOGOROV-SMIRNOV TWO-SAMPLE TEST (large samples: two-tailed test)*

Level of significance	Value of D so large as to call for rejection of H_0 at the indicated level of significance, where D = maximum $\lvert S_{n_1}(X) - S_{n_2}(X)\rvert$
.10	$1.22\sqrt{\frac{n_1+n_2}{n_1n_2}}$
.05	$1.36\sqrt{\frac{n_1+n_2}{n_1n_2}}$
.025	$1.48\sqrt{\frac{n_1+n_2}{n_1n_2}}$
.01	$1.63\sqrt{\frac{n_1+n_2}{n_1n_2}}$
.005	$1.73\sqrt{\frac{n_1+n_2}{n_1n_2}}$
.001	$1.95\sqrt{\frac{n_1+n_2}{n_1n_2}}$

المصدر : نفس المصدر السابق صفحة (٣٥٨) .

جدول (٨)

اختبار فروقات الرتب للأزواج المتقارنة

WILCOXON *T* VALUES

Critical values of *T*, the Wilcoxon signed rank statistic, where *T* is the largest integer such that $Pr(T \leq t/N) \leq \alpha$ the cumulative one-tail probability

	2α .15	.10	05	04	03	.02	01
N	α 075	050	.025	.020	.015	.010	.005
4	0						
5	1	0					
6	2	2	0	0			
7	4	3	2	1	0	0	
8	7	5	3	3	2	1	0
9	9	8	5	5	4	3	1
10	12	10	8	7	6	5	3
11	16	13	10	9	8	7	5
12	19	17	13	12	11	9	7
13	24	21	17	16	14	12	9
14	28	25	21	19	18	15	12
15	33	30	25	23	21	19	15
16	39	35	29	28	26	23	19
17	45	41	34	33	30	27	23
18	51	47	40	38	35	32	27
19	58	53	46	43	41	37	32
20	65	60	52	50	47	43	37
21	73	67	58	56	53	49	42
22	81	75	65	63	59	55	48
23	89	83	73	70	66	62	54
24	98	91	81	78	74	69	61
25	108	100	89	86	82	76	68
26	118	110	98	94	90	84	75
27	128	119	107	103	99	92	83
28	138	130	116	112	108	101	91
29	150	140	126	122	117	110	100
30	161	151	137	132	127	120	109
31	173	163	147	143	137	130	118
32	186	175	159	154	148	140	128
33	199	187	170	165	159	151	138
34	212	200	182	177	171	162	148
35	226	213	195	189	182	173	159
40	302	286	264	257	249	238	220
50	487	466	434	425	413	397	373
60	718	690	648	636	620	600	567
70	995	960	907	891	872	846	805
80	1318	1276	1211	1192	1168	1136	1086
90	1688	1638	1560	1537	1509	1471	1410
100	2105	2045	1955	1928	1894	1850	1779

المصدر :

Patchet (I.S.); Statistical Methods for Managers and ADMINST, VNR, NEWYORK, 1982 page (359).

جدول (٩)

اختبار مجموع الرتب لعينتين

CRITICAL VALUES OF *U* IN THE MANN-WHITNEY TEST

In the first table the entries are the critical values of *U* for a one-tailed test at 0.025 or for a two-tailed test at 0.05; in the second, for a one-tailed test at 0.05 or for a two-tailed test at 0.10.

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																				
2								0	0	0	0	1	1	1	1	1	2	2	2	2
3					0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4				0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5			0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6			1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7			1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8		0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9		0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10		0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11		0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12		1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13		1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14		1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15		1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16		1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17		2	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18		2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19		2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20		2	8	13	20	27	34	41	48	55	62	69	76	83	90	98	105	112	110	127

n_2 \ n_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1																			0	0
2					0	0	0	1	1	1	1	2	2	2	3	3	3	4	4	4
3			0	0	1	2	2	3	3	4	5	5	6	7	7	8	9	9	10	11
4		•	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
5		0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6		0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7		0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
8		1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
9		1	3	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
10		1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
11		1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
12		2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
13		2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
14		2	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
15		3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
16		3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
17		3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
18		4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
19	0	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
20	0	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

المصدر : نفس المصدر السابق صفحة (٣٥٧)

جدول رقم (١٠)

قيم ر الحرجة لاختبار ز = صفراً

للاختبار ذى الجانبين، a قيمتها ضعف القيمة المسجلة عند عنوان العمود الذى له قيمة r الحرجة، لذلك لقيمة 05. = a اختار العمود 025.

α / n	.05	.025	.005
5	.805	.878	.959
6	.729	.811	.917
7	.669	.754	.875
8	.621	.707	.834
9	.582	.666	.798
10	.549	.632	.765
11	.521	.602	.735
12	.497	.576	.708
13	.476	.553	.684
14	.457	.532	.661
15	.441	.514	.641
16	.426	.497	.623

α / n	.05	.025	.005
17	.412	.482	.606
18	.400	.468	.590
19	.389	.456	.575
20	.378	.444	.561
25	.337	.396	.505
30	.306	.361	.463
35	.283	.334	.430
40	.264	.312	.402
50	.235	.279	.361
60	.214	.254	.330
80	.185	.220	.286
100	.165	.196	.256

المصدر : نفس المصدر المذكور فى الفصل العاشر.

جدول (١١)

جدول تحويل ر إلى ى (تحويل فيشر لمعامل الارتباط)

r	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
.0	0.000	0.010	0.020	0.030	0.040	0.050	0.060	0.070	0.080	0.090
.1	.100	.110	.121	.131	.141	.151	.161	.172	.182	.192
.2	.203	.213	.224	.234	.245	.255	.266	.277	.288	.299
.3	.310	.321	.332	.343	.354	.365	.377	.388	.400	.412
.4	.424	.436	.448	.460	.472	.485	.497	.510	.523	.536
.5	.549	.563	.576	.590	.604	.618	.633	.648	.662	.678
.6	.693	.709	.725	.741	.758	.775	.793	.811	.829	.848
.7	.867	.887	.908	.929	.950	.973	.996	1.020	1.045	1.071
.8	1.099	1.127	1.157	1.188	1.221	1.256	1.293	1.333	1.376	1.422
r	0.000	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009
.90	1.472	1.478	1.483	1.488	1.494	1.499	1.505	1.510	1.516	1.522
.91	1.528	1.533	1.539	1.545	1.551	1.557	1.564	1.570	1.576	1.583
.92	1.589	1.596	1.602	1.609	1.616	1.623	1.630	1.637	1.644	1.651
.93	1.658	1.666	1.673	1.681	1.689	1.697	1.705	1.713	1.721	1.730
.94	1.738	1.747	1.756	1.764	1.774	1.783	1.792	1.802	1.812	1.822
.95	1.832	1.842	1.853	1.863	1.874	1.886	1.897	1.909	1.921	1.933
.96	1.946	1.959	1.972	1.986	2.000	2.014	2.029	2.044	2.060	2.076
.97	2.092	2.109	2.127	2.146	2.165	2.185	2.205	2.227	2.249	2.273
.98	2.298	2.323	2.351	2.380	2.410	2.443	2.477	2.515	2.555	2.599
.99	2.646	2.700	2.759	2.826	2.903	2.994	3.106	3.250	3.453	3.800

المصدر : نفس المصدر المذكور، ص (٣٩٨).

جدول (١٢)

دوال لغة بيسك VS BASIC علىإعلى جهاز IBM 3033

Function Name	Purpose	Category
ABS(X)	Absolute Value	Intrinsic
ACOS(X)	Arccosine	Intrinsic
AIDX(A) or AIDX(A$)	Ascending Index	Array
ANGLE(X,Y)	Angle	Intrinsic
ASIN(X)	Arcsine	Intrinsic
ASORT(A) or ASORT(A$)	Ascending Sort	Array
ATN(X)	Arctangent	Intrinsic
CEIL(X)	Ceiling	Intrinsic
CEN(X)	Fahrenheit to Centigrade	Intrinsic
CHR$(M)	Character	Intrinsic
CNT	Count	Intrinsic
CODE	Code	Intrinsic
CON	Constant	Array
COS(X)	Cosine	Intrinsic
COSH(X)	Hyperbolic Cosine	Intrinsic
COT(X)	Cotangent	Intrinsic
CSC(X)	Cosecant	Intrinsic
DAT$[M]	(Year/Month/Day)	Intrinsic
DATE	(Year, Number of Days)	Intrinsic
DATE$	(Year/Month/Day)	Intrinsic
DBL(X)	Real Double	Intrinsic
DEC(X)	Decimal	Intrinsic
DEG(X)	Radians to Degrees	Intrinsic
DET[(A)]	Determinant	Intrinsic
DIDX(A) or DIDX(A$)	Descending Index	Array
DOT(A,B)	Dot Product	Intrinsic
DSORT(A) or DSORT(A$)	Descending Sort	Array
EPS	ε	Intrinsic
ERR	Exception Code	Intrinsic
EXP(X)	Exponential Value	Intrinsic
FAH(X)	Centigrade to Fahrenheit	Intrinsic
FILE(M)	File Status	Intrinsic
FILENUM	File Number	Intrinsic
FILE$(M)	File Name	Intrinsic
FP(X)	Fractional Part	Intrinsic
IDN	Identity	Array
IFIX(X)	Rounded Integer Value	Intrinsic
INF	Infinity	Intrinsic
INT(X)	Largest Integer	Intrinsic
INV(A)	Inverse	Array
IP(X)	Integer Part	Intrinsic
JDY[(C$)]	Julian Date	Intrinsic
KEYNUM	Key Number	Intrinsic
KLN(M)	Key Length	Intrinsic

المصدر : IBM 3270 Bashanced Functions TEXT Book

تابع جدول (١٢)

Function Name	Purpose	Category
KPS(M)	Key Position	Intrinsic
LEN(C$)	Length	Intrinsic
LINE	Line Number	Intrinsic
LOG(X)	Natural Logarithm	Intrinsic
LOG2(X)	Base 2 Logarithm	Intrinsic
LOG10(X)	Common Logarithm	Intrinsic
LPAD$(C$,M)	Left Pad	Intrinsic
LTRM$(C$)	Left Trim	Intrinsic
LWRC$(C$)	Lower Case	Intrinsic
MAX(X,Y[,Z]...)	Maximum	Intrinsic
MIN(X,Y[,Z]...)	Minimum	Intrinsic
MOD(X,Y)	Modulo	Intrinsic
NUL$	Null String	Array
ORD(C$)	Ordinal Position	Intrinsic
PARM$	Parameter	Intrinsic
PI	π	Intrinsic
POS(C$,D$)	Position	Intrinsic
POS(C$,D$,M)	Position	Intrinsic
PRD(A)	Array Product	Intrinsic
RAD(X)	Degrees to Radians	Intrinsic
REAL(X)	Real	Intrinsic
REC(M)	Record Number	Intrinsic
REM(X,Y)	Remainder	Intrinsic
RLN(M)	Record Length	Intrinsic
RND[(X)]	Random	Intrinsic
ROUND(X,M)	Round	Intrinsic
RPAD$(C$,M)	Right Pad	Intrinsic
RPT$(C$,M)	Repeat	Intrinsic
RTRM$(C$)	Right Trim	Intrinsic
SEC(X)	Secant	Intrinsic
SGN(X)	Sign	Intrinsic
SIN(X)	Sine	Intrinsic
SINH(X)	Hyperbolic Sine	Intrinsic
SIZE(A) or SIZE(A$)	Array Size	Intrinsic
SIZE(A,M) or SIZE(A$,M)	Dimension Size	Intrinsic
SNG(X)	Real Single	Intrinsic
SQR(X)	Square Root	Intrinsic
SRCH(A,X[,M])	Numeric Search	Intrinsic
SRCH(A$,C$[,M])	Character Search	Intrinsic
SREP$(C$,M,D$,E$)	Search and Replace	Intrinsic
STR$(X)	String Conversion	Intrinsic
SUM(A)	Sum	Intrinsic
TAN(X)	Tangent	Intrinsic
TANH(X)	Hyperbolic Tangent	Intrinsic
TIME	Time in Seconds	Intrinsic
TIME$	Time (HH:MM:SS)	Intrinsic
TRN(A) or TRN(A$)	Transpose	Array
TRUNCATE(X,M)	Truncate	Intrinsic
UDIM(A,M) or UDIM(A$,M)	Upper Dimension	Intrinsic



- ١ ▪ المراجـع العربية
- ٢ ▪ التقارير العربية
- ٣ ▪ المراجع الأجنبية



١ ـ المراجع العربية:

١ - بول ٠ ج ٠ هويل : المبادىء الأولية فى الإحصاء - ترجمة د٠ بدرية عبدالوهاب و د٠ محمد الشربينى - الطبعة الرابعة - جون وايلى وأبنائه - ١٩٨٤م (نيويورك) .

٢ - د٠ على عبدالحفيظ : دور وحدات التخطيط فى الأجهزة الحكومية فى المملكة العربية السعودية - معهد الإدارة العامة - الرياض - ١٤٠٤هـ .

٣ - د٠ فاروق عبدالعظيم أحمد : مقدمة الطرق الإحصائية - دار المطبوعات الجامعية - الاسكندرية - ١٩٧٩ .

٤ - محمد صبحى أبو صالح وعدنان محمد عوض : مقدمة فى الإحصاء ، جون وايلى وأبنـــائه - نيويورك ١٩٨٣ .

٥ ـ د. محمد مظلوم حمدى : طرق الإحصاء ـ دار المعارف بمصر ـ الطبعة الرابعة ـ مصر ـ ١٩٦١ .

٢ ـ التقارير العربية:

١ ـ بحث حوادث السيارات والاضرار الصحية الناتجة عنها ـ الادارة العامة للمرور ـ الرياض ١٩٨١ .

٢ ـ النشرة الاحصائية السنوية لحوادث المرور ـ الادارة العامة للمرور ـ الرياض ـ ١٤٠٢هـ .

٣ ـ مؤسسة النقد السعودى ـ التقرير السنوى لعام ١٤٠١هـ ـ الرياض ١٤٠٢هـ .

٤ ـ مؤسسة النقد السعودى ـ التقرير السنوى لعام ١٤٠٤هـ ـ الرياض ١٤٠٥هـ.

٣ ـ المراجع الأجنبية:

1 - Achen (C. H.); ***Interpreting and using Regression,*** Sage pubilcations, London, 1982.

2 - Afifi (A.A.) and Azen (S.P.); ***Statistical Analysis, A computer Approach,*** Academic press Inc., New York, 1972.

3 - Barrie (W.G.); ***Intermediate Statistical Methods,*** Chapman and Hall, London, 1981.

4 - Clark (F.J.); ***Mathematics for Data Processing,*** Reston Publishing Co., 1983.

5 - Dixon (W.J.) and Massey (F.J.); ***Introduction to Statistical Analysis,*** Fourth edition, Mc-Graw. Hill, New York, 1983.

6 - Dormey (R.G.); ***How to Solve it by Computer,*** Prentice-Hall, 1983.

7 - Draper (N.R.) and Smith (H.); ***Applied Regression Analysis,*** John wiley, New York, 1966.

8 – Duncan (A.J.); ***Quality Control and Industrial Statistics,*** Irwin, London, 1973.

9 – Fadil (H.Z.); ***Applied Business Statistics,*** Addison – Wesley, Reading, U.S.A, 1984.

10 – Freund (J.E.); ***Mathematical Statistics,*** Hall International, London, fourth edition, 1972.

11 – Freuwd (R.J.); ***Regression Methods,*** Marcel Dekker, New York, 1980.

12 – Horowitz (E.) and Sahni (S.); ***Fundamentals of Computer Algorithms,*** Pitman Pub. Co., 1978.

13 – Jones (R.M.); ***Structured Basic,*** Allyn and Bacon, 1985.

14 – Kazemeir (L.J.); ***Statistical Analysis for Business and Economics,*** Mc-Graw – Hill, New York; Third edition, 1978.

15 – Kendall (M.) and Stuart (A.); ***Advanced Theory of Statistics,*** Vol. 2, Griffin; London; Third edition, 1961.

16 – Kitchen (A.); ***Basic By Decision,*** Prentice-Hall, 1983.

17 – Larson (H.J.); ***Introduction To Probability Theory and Statistical Inference,*** John Wiley, NewYork, Second edition,1974.

18 – Lien (D.A.); ***THE Basic Handbook,*** Compuso Publishing, 1981.

19 – Lipschutz (S.); ***Essential computer Mathematics,*** Mc-Graw Hall, Company, 1982.

20 – Mason (R.D.); ***Statistical Techniques in Businen and Economics,*** Richard D. Irwin, Illinois, Third edition, 1974.

21 – Meier (K.J.) and Brudney (J.L.); ***Applied Statistics for public Administration,*** Duxbury Press, U.S.A., 1981.

22 – Michael (S. Lewis-Beck); ***Applied Regression, an Introduction,*** Sage University Paper 22, London, Fifth Printing, 1983.

23 – Meyer (P.L.); ***Introduction to Probability and Statistical Applications,*** Addison-Wesley, California, Second edition, 1972.

24 – Patchet (I.S.); ***Statistical Methods for Managers and Administrators,*** Van Nostr. Reinhold Company, New York, 1980.

25 – Poole (C.); Borche (M.) and Castle (D.); ***Some Common Basic Programs,*** McGraw-Hill, 1981.

26 – Roberts (H.V.) and Robert (F.L.); ***An Introduction to data Analysis and Regression,*** McGraw-Hill, New York, 1982.

27 – Ronald (S.K.) and Bryant (J.); ***Applied Statisitics Using The Computer,*** Alfred Pub. Co., California, 1982.

28 – Snedecor (G.W.) and Cochran (W.); ***Statistical Methods,*** Iowa State University Press, IWA, Sixth Edition, 7th Printing, 1974.

29 – Yeomans (P.S.); ***Applied Statistics for Social Scientist,*** Volume two, Penguin Books, Middlesex, England.

30 – Waller (R. A.); ***Statistics: An Intoduction to Numerical Reasoning,*** Holland Inc; San Francisco, 1979.

المؤلفان فى سطور

●● الأستاذ محمد عثمان البشير.

- من مواليد الدويم - السودان.
- حاصل على درجة الماجستير فى مجال الحاسب الآلى من معهد شمال لندن التكنولوجى فى عام ١٩٧٨.
- يعمل حاليًا رئيسًا لقسم الحاسب الآلى بفرع المعهد بالغربية.

● من خبراته العملية :

- محاضر ومحلل نظم / مبرمج مركز الحاسب الآلى، جامعة الخرطوم، محاضر ومحلل نظم / مبرمج بمعهد الإدارة العامة، رئيس قسم الحاسب الآلى بفرع معهد الإدارة العامة بالغربية.

● من اهم اعماله العلمية المنشورة :

- مقدمة فى الحاسب الآلى (كتاب).
- الكمبيوتر الشخصى ما هو؟ (مقال نشر فى مجلة الإدارة العامة).

●● الأستاذ كرم الله على عبدالرحمن.

- من مواليد السودان.
- حصل على درجة الماجستير فى الإحصاء (إحصاء تطبيقى) من جامعة ساوثهامبتون بإنجلترا عام ١٩٧٧م، ويعمل حاليًا محاضرًا بمعهد الإدارة العامة.

● من خبراته العملية :

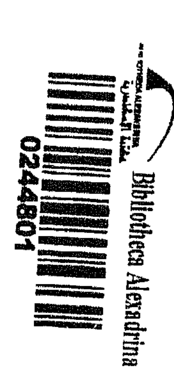
- إخصائى بقسم الاقتصاد والإحصاء الزراعى بالسودان، باحث بالمجلس القومى للبحوث بالسودان، محاضر بجامعة الخرطوم، وأخيرًا محاضر بمعهد الإدارة العامة.

● من اهم اعماله العلمية المنشورة :

- The Robustness of Trimmed Estimators in Linear Regression (بحث).
- حوادث المرور (بحث).
- الرغبة فى التسرب لدى خريجى البرامج الإعدادية (بحث).
- عدد من المقالات فى مجال الإحصاء نشرت بمجلة الإدارة العامة.

طبعت بمطابع معهد الإدارة العامة ١٤١١هـ





طبعت بمطابع معهد الإدارة العامة ١٤١١هـ

٣٦ ريالاً